

I

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
26 June 2003 (26.06.2003)

PCT

(10) International Publication Number
WO 03/052052 A2

(51) International Patent Classification⁷: C12N

(74) Agents: **KODROFF, Cathy, A.** et al.: Howson and Howson, Spring House Corporate Center, P.O. Box 457, Spring House, PA 19477 (US).

(21) International Application Number: PCT/US02/33631

(22) International Filing Date:
12 November 2002 (12.11.2002)

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/341,150 17 December 2001 (17.12.2001) US
60/386,132 5 June 2002 (05.06.2002) US

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): **THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA** [US/US]; 3160 Chestnut Street, Suite 200, Philadelphia, PA 19104 (US).

Published:

— without international search report and to be republished upon receipt of that report

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **GAO, Guangping** [US/US]; 408 Yorkshire Road, Rosemont, PA 19010 (US). **WILSON, James, M.** [US/US]; 1350 N. Avignon Drive, Gladwyne, PA 19035 (US). **ALVIRA, Mauricio** [US/US]; 4417 Pine Street, Apt. 203, Philadelphia, PA 19104 (US).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ADENO-ASSOCIATED VIRUS (AAV) SEROTYPE 9 SEQUENCES, VECTORS CONTAINING SAME, AND USES THEREFOR

(57) Abstract: Sequences of a serotype 9 adeno-associated virus and vectors and host cells containing these sequences are provided. Also described are methods of using such host cells and vectors in production of rAAV particles. RAAV9-mediated delivery of therapeutic and immunogenic genes is also provided.



WO 03/052052 A2

ADENO-ASSOCIATED VIRUS (AAV) SEROTYPE 9 SEQUENCES, VECTORS CONTAINING SAME, AND USES THEREFOR

BACKGROUND OF THE INVENTION

- 5 Adeno-associated virus (AAV), a member of the Parvovirus family, is a small nonenveloped, icosahedral virus with single-stranded linear DNA genomes of 4.7 kilobases (kb) to 6 kb. AAV is assigned to the genus, *Dependovirus*, because the virus was discovered as a contaminant in purified adenovirus stocks. AAV's life cycle includes a latent phase at which AAV genomes, after infection, are site specifically
- 10 integrated into host chromosomes and an infectious phase in which, following either adenovirus or herpes simplex virus infection, the integrated genomes are subsequently rescued, replicated, and packaged into infectious viruses. The properties of non-pathogenicity, broad host range of infectivity, including non-dividing cells, and potential site-specific chromosomal integration make AAV an attractive tool for gene transfer.
- 15 Recent studies suggest that AAV vectors may be the preferred vehicle for gene delivery. To date, there have been 6 different serotypes of AAVs isolated from human or non-human primates (NHP) and well characterized. Among them, human serotype 2 is the first AAV that was developed as a gene transfer vector; it has been widely used for efficient gene transfer experiments in different target tissues and animal models.
- 20 Clinical trials of the experimental application of AAV2 based vectors to some human disease models are in progress, and include such diseases as cystic fibrosis and hemophilia B.

What are desirable are AAV-based constructs for gene delivery.

25 SUMMARY OF THE INVENTION

- In one aspect, the invention provides novel AAV sequences, compositions containing these sequences, and uses therefor. Advantageously, these compositions are particularly well suited for use in compositions requiring re-administration of rAAV for therapeutic or prophylactic purposes.
- 30 These and other aspects of the invention will be readily apparent from the following detailed description of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1A, 1B and 1C are the nucleic acid sequences of the rep and cap regions of AAV9 [SEQ ID NO:1].

5 Figs. 2A, 2B and 2C are an alignment of the amino acid sequences of the vp1 protein of the capsid of the novel AAV9 [SEQ ID NO:2] sequences of the invention as compared to the published sequences of AAV2 [SEQ ID NO:4], AAV1 [SEQ ID NO:5], and AAV3 [SEQ ID NO:6], and of a novel serotype AAV8 [SEQ ID NO:7], which is the subject of a co-pending application. The alignment was performed using the Clustal W program, with the numbering of AAV2 used for reference. Underlining and bold under
10 the AAV9 sequences indicates cassettes of identity within the HVR.

Figs. 3A, 3B and 3C are the amino acid sequences of the rep proteins of AAV9 [SEQ ID NO:3].

DETAILED DESCRIPTION OF THE INVENTION

15 The invention provides the nucleic acid sequences and amino acids of a novel AAV serotype, AAV9. Also provided are fragments of these AAV sequences. Each of these fragments may be readily utilized in a variety of vector systems and host cells. Among desirable AAV9 fragments are the cap proteins, including the vp1, vp2, vp3 and hypervariable regions, the rep proteins, including rep 78, rep 68, rep 52, and rep
20 40, and the sequences encoding these proteins. These fragments may be readily utilized in a variety of vector systems and host cells. Such fragments may be used alone, in combination with other AAV9 sequences or fragments, or in combination with elements from other AAV or non-AAV viral sequences. In one particularly desirable embodiment, a vector contains the AAV9 cap and/or rep sequences of the invention.

25 The AAV9 sequences and fragments thereof are useful in production of rAAV, and are also useful as antisense delivery vectors, gene therapy vectors, or vaccine vectors. The invention further provides nucleic acid molecules, gene delivery vectors, and host cells which contain the AAV9 sequences of the invention.

Suitable fragments can be determined using the information provided herein.
30 Alignments are performed using any of a variety of publicly or commercially available Multiple Sequence Alignment Programs, such as "Clustal W", accessible through Web Servers on the internet. Alternatively, Vector NTI utilities are also used. There are also a

number of algorithms known in the art which can be used to measure nucleotide sequence identity, including those contained in the programs described above. As another example, polynucleotide sequences can be compared using Fasta, a program in GCG Version 6.1. Fasta provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences. For instance, percent sequence identity between nucleic acid sequences can be determined using Fasta with its default parameters (a word size of 6 and the NOPAM factor for the scoring matrix) as provided in GCG Version 6.1, herein incorporated by reference. Similar programs are available for amino acid sequences, e.g., the "Clustal X" program.

Generally, any of these programs are used at default settings, although one of skill in the art can alter these settings as needed. Alternatively, one of skill in the art can utilize another algorithm or computer program which provides at least the level of identity or alignment as that provided by the referenced algorithms and programs.

The term "substantial homology" or "substantial similarity," when referring to a nucleic acid, or fragment thereof, indicates that, when optimally aligned with appropriate nucleotide insertions or deletions with another nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 95 to 99% of the aligned sequences. Preferably, the homology is over full-length sequence, or an open reading frame thereof, or another suitable fragment which is at least 15 nucleotides in length. Examples of suitable fragments are described herein.

The term "substantial homology" or "substantial similarity," when referring to amino acids or fragments thereof, indicates that, when optimally aligned with appropriate amino acid insertions or deletions with another amino acid (or its complementary strand), there is amino acid sequence identity in at least about 95 to 99% of the aligned sequences. Preferably, the homology is over full-length sequence, or a protein thereof, e.g., a cap protein, a rep protein, or a fragment thereof which is at least 8 amino acids, or more desirably, at least 15 amino acids in length. Examples of suitable fragments are described herein.

By the term "highly conserved" is meant at least 80% identity, preferably at least 90% identity, and more preferably, over 97% identity. Identity is readily determined by one of skill in the art by resort to algorithms and computer programs known by those of skill in the art.

The term “percent sequence identity” or “identical” in the context of nucleic acid sequences refers to the residues in the two sequences which are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over the full-length of the genome, the full-length of a gene coding sequence, or a fragment of at least about 500 to 5000 nucleotides, is desired. However, identity among smaller fragments, e.g. of at least about nine nucleotides, usually at least about 20 to 24 nucleotides, at least about 28 to 32 nucleotides, at least about 36 or more nucleotides, may also be desired. Similarly, “percent sequence identity” may be readily determined for amino acid sequences, over the full-length of a protein, or a fragment thereof. Suitably, a fragment is at least about 8 amino acids in length, and may be up to about 700 amino acids. Examples of suitable fragments are described herein.

As described herein, the vectors of the invention containing the AAV capsid proteins of the invention are particularly well suited for use in applications in which the neutralizing antibodies diminish the effectiveness of other AAV serotype based vectors, as well as other viral vectors. The rAAV vectors of the invention are particularly advantageous in rAAV readministration and repeat gene therapy.

These and other embodiments and advantages of the invention are described in more detail below. As used throughout this specification and the claims, the term “comprising” is inclusive of other components, elements, integers, steps and the like. Conversely, the term “consisting” and its variants are exclusive of other components, elements, integers, steps and the like.

I. AAV Serotype 9 Sequences

A. Nucleic Acid Sequences

The AAV9 nucleic acid sequences of the invention include the DNA sequences of Fig. 1 [SEQ ID NO: 1], which consists of 4382 nucleotides. The AAV9 nucleic acid sequences of the invention further encompass the strand which is complementary to Fig. 1 [SEQ ID NO: 1], as well as the RNA and cDNA sequences corresponding to Fig. 1 [SEQ ID NO: 1] and its complementary strand. Also included in the nucleic acid sequences of the invention are natural variants and engineered modifications of Fig. 1 [SEQ ID NO: 1] and its complementary strand. Such

modifications include, for example, labels which are known in the art, methylation, and substitution of one or more of the naturally occurring nucleotides with a degenerate nucleotide.

Further included in this invention are nucleic acid sequences which are greater than 85%, preferably at least about 90%, more preferably at least about 95%, and most preferably at least about 98 to 99% identical or homologous to Fig. 1 [SEQ ID NO:1]. Also included within the invention are fragments of Fig. 1 [SEQ ID NO: 1], its complementary strand, cDNA and RNA complementary thereto. Suitable fragments are at least 15 nucleotides in length, and encompass functional fragments, i.e., fragments which are of biological interest. Such fragments include the sequences encoding the three variable proteins (vp) of the AAV9 capsid which are alternative splice variants: vp1 [nt 2116 to 4323 of Fig. 1, SEQ ID NO:1]; vp2 [nt 2527 to 4323 of Fig. 1, SEQ ID NO:1]; and vp 3 [nt 2725 to 4323 of Fig. 1, SEQ ID NO:1]. Other suitable fragments of Fig. 1, SEQ ID NO:1, the fragment which contains the start codon for the AAV9 capsid protein.

Still other fragments include those encoding the rep proteins, including *rep* 78 [initiation codon at nt 228 of Fig. 1], *rep* 68 [initiation codon at nt 228 of Fig. 1], *rep* 52 [initiation codon at nt 900 of Fig. 1], and *rep* 40 [initiation codon at nt 900 of Fig. 1]. See, SEQ ID NO:1. Other fragments of interest may include the AAV inverted terminal repeats (ITRs), AAV P19 sequences, AAV P40 sequences, the rep binding site, and the terminal resolute site (TRS). Still other suitable fragments will be readily apparent to those of skill in the art.

In addition to including the nucleic acid sequences provided in the figures and Sequence Listing, the present invention includes nucleic acid molecules and sequences which are designed to express the amino acid sequences, proteins and peptides of the AAV serotypes of the invention. Thus, the invention includes nucleic acid sequences which encode the following novel AAV amino acid sequences and artificial AAV serotypes generated using these sequences and/or unique fragments thereof.

As used herein, artificial AAV serotypes include, without limitation, AAV with a non-naturally occurring capsid protein. Such an artificial capsid may be generated by any suitable technique, using a novel AAV sequence of the invention (e.g., a fragment of a vp1 capsid protein) in combination with heterologous sequences which may be obtained from another AAV serotype (known or novel), non-contiguous portions

of the same AAV serotype, from a non-AAV viral source, or from a non-viral source. An artificial AAV serotype may be, without limitation, a chimeric AAV capsid, a recombinant AAV capsid, or a "humanized" AAV capsid.

B. AAV9 Amino Acid Sequences, Proteins and Peptides

5 The invention further provides proteins and fragments thereof which are encoded by the AAV9 nucleic acids of the invention, and AAV9 amino acids which are generated by other methods. The invention further encompasses AAV serotypes generated using sequences of the novel AAV serotype of the invention, which are generated using synthetic, recombinant or other techniques known to those of skill in the art. The invention is not limited to novel AAV amino acid sequences, peptides and
10 proteins expressed from the novel AAV nucleic acid sequences of the invention and encompasses amino acid sequences, peptides and proteins generated by other methods known in the art, including, e.g., by chemical synthesis, by other synthetic techniques, or by other methods. For example, the sequences of any of be readily generated using a
15 variety of techniques.

Suitable production techniques are well known to those of skill in the art. See, e.g., Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (Cold Spring Harbor, NY). Alternatively, peptides can also be synthesized by the well known solid phase peptide synthesis methods (Merrifield, *J. Am. Chem. Soc.*,
20 85:2149 (1962); Stewart and Young, *Solid Phase Peptide Synthesis* (Freeman, San Francisco, 1969) pp. 27-62). These and other suitable production methods are within the knowledge of those of skill in the art and are not a limitation of the present invention.

Particularly desirable proteins include the AAV9 capsid proteins and AAV9 rep proteins.

25 Particularly desirable proteins include the AAV capsid proteins, which are encoded by the nucleotide sequences identified above. The AAV capsid is composed of three proteins, vp1, vp2 and vp3, which are alternative splice variants. The full-length sequence provided in figure 2 is that of vp1. The AAV9 capsid proteins include vp1 [SEQ ID NO:2], vp2 [aa 138 to 736 of SEQ ID NO:2], and vp3 [aa 203 to 736 of SEQ ID
30 NO: 2] and functional fragments thereof. Other desirable fragments of the capsid protein include the constant and variable regions, located between hypervariable regions (HPV). Other desirable fragments of the capsid protein include the HPV themselves.

An algorithm developed to determine areas of sequence divergence in AAV2 has yielded 12 hypervariable regions (HVR) of which 5 overlap or are part of the four previously described variable regions. [Chiorini *et al*, *J. Virol*, 73:1309-19 (1999); Rutledge *et al*, *J. Virol.*, 72:309-319] Using this algorithm and/or the alignment techniques described herein, the HVR of the novel AAV serotypes are determined. For example, with respect to the number of the AAV2 vp1 [SEQ ID NO:4], the HVR are located as follows: HVR1, aa 146-152; HVR2, aa 182-186; HVR3, aa 262-264; HVR4, aa 381-383; HVR5, aa 450-474; HVR6, aa 490-495; HVR7, aa500-504; HVR8, aa 514-522; HVR9, aa 534-555; HVR10, aa 581-594; HVR11, aa 658-667; and HVR12, aa 705-719. Using the alignment provided herein performed using the Clustal X program at default settings, or using other commercially or publicly available alignment programs at default settings, one of skill in the art can readily determine corresponding fragments of the novel AAV capsids of the invention.

Still other desirable fragments of the AAV9 capsid protein include amino acids 1 to 184 of SEQ ID NO: 2, amino acids 199 to 259; amino acids 274 to 446; amino acids 603 to 659; amino acids 670 to 706; amino acids 724 to 736 of SEQ ID NO:2; aa 185 - 198; aa 260-273; aa447-477; aa495-602; aa660-669; and aa707-723. Additionally, examples of other suitable fragments of AAV capsids include, with respect to the numbering of AAV2 [SEQ ID NO:4], aa 24 to 42, aa 25 to 28; aa 81 to 85; aa133 to 165; aa 134 to 165; aa 137 to 143; aa 154 to 156; aa 194 to 208; aa 261 to 274; aa 262 to 274; aa 171 to 173; aa 413 to 417; aa 449 to 478; aa 494 to 525; aa 534 to 571; aa 581 to 601; aa 660 to 671; aa 709 to 723. Still other desirable fragments include, for example, in AAV7, amino acids 1 to 184 of SEQ ID NO:2, amino acids 199 to 259; amino acids 274 to 446; amino acids 603 to 659; amino acids 670 to 706; amino acids 724 to 736; aa 185 to 198; aa 260 to 273; aa447 to 477; aa495 to 602; aa660 to 669; and aa707 to 723. Using the alignment provided herein performed using the Clustal X program at default settings, or using other commercially or publicly available alignment programs at default settings, one of skill in the art can readily determine corresponding fragments of the novel AAV capsids of the invention.

Still other desirable AAV9 proteins include the rep proteins include rep68/78 and rep40/52 [located within SEQ ID NO: 3]. Suitable fragments of the rep proteins may include aa 1 to 102; aa 103 to 140; aa 141 to 173; aa 174 to 226; aa 227 to

275; aa 276 to 374; aa 375 to 383; aa 384 to 446; aa 447 to 542; aa 543 to 555; aa 556 to 623, of SEQ ID NO: 3.

Suitably, fragments are at least 8 amino acids in length. However, fragments of other desired lengths may be readily utilized. Such fragments may be produced recombinantly or by other suitable means, e.g., chemical synthesis.

The invention further provides other AAV9 sequences which are identified using the sequence information provided herein. For example, given the AAV9 sequences provided herein, infectious AAV9 may be isolated using genome walking technology (Siebert *et al.*, 1995, *Nucleic Acid Research*, 23:1087-1088, Frieznner-Degen *et al.*, 1986, *J. Biol. Chem.* 261:6972-6985, BD Biosciences Clontech, Palo Alto, CA). Genome walking is particularly well suited for identifying and isolating the sequences adjacent to the novel sequences identified according to the method of the invention. This technique is also useful for isolating inverted terminal repeat (ITRs) of the novel AAV9 serotype, based upon the novel AAV capsid and rep sequences provided herein.

The sequences, proteins, and fragments of the invention may be produced by any suitable means, including recombinant production, chemical synthesis, or other synthetic means. Such production methods are within the knowledge of those of skill in the art and are not a limitation of the present invention.

IV. Production of rAAV with AAV9 Capsids

The invention encompasses novel, wild-type AAV9, the sequences of which are free of DNA and/or cellular material with these viruses are associated in nature. In another aspect, the present invention provides molecules which utilize the novel AAV sequences of the invention, including fragments thereof, for production of molecules useful in delivery of a heterologous gene or other nucleic acid sequences to a target cell.

In another aspect, the present invention provides molecules which utilize the AAV9 sequences of the invention, including fragments thereof, for production of viral vectors useful in delivery of a heterologous gene or other nucleic acid sequences to a target cell.

The molecules of the invention which contain AAV9 sequences include any genetic element (vector) which may be delivered to a host cell, e.g., naked DNA, a plasmid, phage, transposon, cosmid, episome, a protein in a non-viral delivery vehicle

(e.g., a lipid-based carrier), virus, etc. which transfer the sequences carried thereon. The selected vector may be delivered by any suitable method, including transfection, electroporation, liposome delivery, membrane fusion techniques, high velocity DNA-coated pellets, viral infection and protoplast fusion. The methods used to construct any embodiment of this invention are known to those with skill in nucleic acid manipulation and include genetic engineering, recombinant engineering, and synthetic techniques. See, e.g., Sambrook et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY.

In one embodiment, the vectors of the invention contain, at a minimum, sequences encoding an AAV9 capsid or a fragment thereof. In another embodiment, the vectors of the invention contain, at a minimum, sequences encoding an AAV9 rep protein or a fragment thereof. Optionally, such vectors may contain both AAV cap and rep proteins. In vectors in which both AAV *rep* and *cap* are provided, the AAV *rep* and AAV *cap* sequences can both be of AAV9 origin. Alternatively, the present invention provides vectors in which the *rep* sequences are from an AAV serotype which differs from that which is providing the *cap* sequences. In one embodiment, the *rep* and *cap* sequences are expressed from separate sources (e.g., separate vectors, or a host cell and a vector). In another embodiment, these *rep* sequences are fused in frame to *cap* sequences of a different AAV serotype to form a chimeric AAV vector. Optionally, the vectors of the invention further contain a minigene comprising a selected transgene which is flanked by AAV 5' ITR and AAV 3' ITR.

Thus, in one embodiment, the vectors described herein contain nucleic acid sequences encoding an intact AAV capsid which may be from a single AAV serotype (e.g., AAV9). Such a capsid may comprise amino acids 1 to 736 of SEQ ID NO:2. Alternatively, these vectors contain sequences encoding artificial capsids which contain one or more fragments of the AAV9 capsid fused to heterologous AAV or non-AAV capsid proteins (or fragments thereof). These artificial capsid proteins are selected from non-contiguous portions of the AAV9 capsid or from capsids of other AAV serotypes. For example, a rAAV may have a capsid protein comprising one or more of the AAV9 capsid regions selected from the vp2 and/or vp3, or from vp 1, or fragments thereof selected from amino acids 1 to 184, amino acids 199 to 259; amino acids 274 to 446; amino acids 603 to 659; amino acids 670 to 706; amino acids 724 to 736 of the AAV9

capsid, SEQ ID NO: 2. In another example, it may be desirable to alter the start codon of the vp3 protein to GTG. Alternatively, the rAAV may contain one or more of the AAV serotype 9 capsid protein hypervariable regions which are identified herein, or other fragment including, without limitation, aa 185 to 198; aa 260-273; aa 447 to 477; aa 495 to 602; aa 660 to 669; and aa 707 to 723 of the AAV9 capsid. See, SEQ ID NO: 2. These modifications may be to increase expression, yield, and/or to improve purification in the selected expression systems, or for another desired purpose (e.g., to change tropism or alter neutralizing antibody epitopes).

The vectors described herein, e.g., a plasmid, are useful for a variety of purposes, but are particularly well suited for use in production of a rAAV containing a capsid comprising AAV sequences or a fragment thereof. These vectors, including rAAV, their elements, construction, and uses are described in detail herein.

In one aspect, the invention provides a method of generating a recombinant adeno-associated virus (AAV) having an AAV serotype 9 capsid, or a portion thereof. Such a method involves culturing a host cell which contains a nucleic acid sequence encoding an adeno-associated virus (AAV) serotype 9 capsid protein, or fragment thereof, as defined herein; a functional rep gene; a minigene composed of, at a minimum, AAV inverted terminal repeats (ITRs) and a transgene; and sufficient helper functions to permit packaging of the minigene into the AAV9 capsid protein.

The components required to be cultured in the host cell to package an AAV minigene in an AAV capsid may be provided to the host cell in *trans*. Alternatively, any one or more of the required components (e.g., minigene, *rep* sequences, *cap* sequences, and/or helper functions) may be provided by a stable host cell which has been engineered to contain one or more of the required components using methods known to those of skill in the art. Most suitably, such a stable host cell will contain the required component(s) under the control of an inducible promoter. However, the required component(s) may be under the control of a constitutive promoter. Examples of suitable inducible and constitutive promoters are provided herein, in the discussion of regulatory elements suitable for use with the transgene. In still another alternative, a selected stable host cell may contain selected component(s) under the control of a constitutive promoter and other selected component(s) under the control of one or more inducible promoters. For example, a stable host cell may be generated which is derived from 293 cells (which

contain E1 helper functions under the control of a constitutive promoter), but which contains the rep and/or cap proteins under the control of inducible promoters. Still other stable host cells may be generated by one of skill in the art.

The minigene, *rep* sequences, *cap* sequences, and helper functions required for producing the rAAV of the invention may be delivered to the packaging host cell in the form of any genetic element which transfer the sequences carried thereon. The selected genetic element may be delivered by any suitable method, including those described herein. The methods used to construct any embodiment of this invention are known to those with skill in nucleic acid manipulation and include genetic engineering, recombinant engineering, and synthetic techniques. See, e.g., Sambrook et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY. Similarly, methods of generating rAAV virions are well known and the selection of a suitable method is not a limitation on the present invention. See, e.g., K. Fisher et al, *J. Virol.*, 70:520-532 (1993) and US Patent 5,478,745.

Unless otherwise specified, the AAV ITRs, and other selected AAV components described herein, may be readily selected from among any AAV serotype, including without limitation, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, and the novel serotype of the invention, AAV9. These ITRs or other AAV components may be readily isolated using techniques available to those of skill in the art from an AAV serotype.

Such AAV may be isolated or obtained from academic, commercial, or public sources (e.g., the American Type Culture Collection, Manassas, VA). Alternatively, the AAV sequences may be obtained through synthetic or other suitable means by reference to published sequences such as are available in the literature or in databases such as, e.g., GenBank, PubMed, or the like.

A. The Minigene

The minigene is composed of, at a minimum, a transgene and its regulatory sequences, and 5' and 3' AAV inverted terminal repeats (ITRs). In one desirable embodiment, the ITRs of AAV serotype 2 are used. However, ITRs from other suitable serotypes may be selected. It is this minigene which is packaged into a capsid protein and delivered to a selected host cell.

1. The transgene

The transgene is a nucleic acid sequence, heterologous to the vector sequences flanking the transgene, which encodes a polypeptide, protein, or other product of interest. The nucleic acid coding sequence is operatively linked to regulatory components in a manner which permits transgene transcription, translation, and/or expression in a host cell.

The composition of the transgene sequence will depend upon the use to which the resulting vector will be put. For example, one type of transgene sequence includes a reporter sequence, which upon expression produces a detectable signal. Such reporter sequences include, without limitation, DNA sequences encoding β -lactamase, β -galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase, membrane bound proteins including, for example, CD2, CD4, CD8, the influenza hemagglutinin protein, and others well known in the art, to which high affinity antibodies directed thereto exist or can be produced by conventional means, and fusion proteins comprising a membrane bound protein appropriately fused to an antigen tag domain from, among others, hemagglutinin or Myc.

These coding sequences, when associated with regulatory elements which drive their expression, provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence or other spectrographic assays, fluorescent activating cell sorting assays and immunological assays, including enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and immunohistochemistry. For example, where the marker sequence is the LacZ gene, the presence of the vector carrying the signal is detected by assays for beta-galactosidase activity. Where the transgene is green fluorescent protein or luciferase, the vector carrying the signal may be measured visually by color or light production in a luminometer.

However, desirably, the transgene is a non-marker sequence encoding a product which is useful in biology and medicine, such as proteins, peptides, RNA, enzymes, or catalytic RNAs. Desirable RNA molecules include tRNA, dsRNA, ribosomal RNA, catalytic RNAs, and antisense RNAs. One example of a useful

RNA sequence is a sequence which extinguishes expression of a targeted nucleic acid sequence in the treated animal.

The transgene may be used to correct or ameliorate gene deficiencies, which may include deficiencies in which normal genes are expressed at less than normal levels or deficiencies in which the functional gene product is not expressed. A preferred type of transgene sequence encodes a therapeutic protein or polypeptide which is expressed in a host cell. The invention further includes using multiple transgenes, e.g., to correct or ameliorate a gene defect caused by a multi-subunit protein. In certain situations, a different transgene may be used to encode each subunit of a protein, or to encode different peptides or proteins. This is desirable when the size of the DNA encoding the protein subunit is large, e.g., for an immunoglobulin, the platelet-derived growth factor, or a dystrophin protein. In order for the cell to produce the multi-subunit protein, a cell is infected with the recombinant virus containing each of the different subunits. Alternatively, different subunits of a protein may be encoded by the same transgene. In this case, a single transgene includes the DNA encoding each of the subunits, with the DNA for each subunit separated by an internal ribozyme entry site (IRES). This is desirable when the size of the DNA encoding each of the subunits is small, e.g., the total size of the DNA encoding the subunits and the IRES is less than five kilobases. As an alternative to an IRES, the DNA may be separated by sequences encoding a 2A peptide, which self-cleaves in a post-translational event. See, e.g., M.L. Donnelly, *et al*, *J. Gen. Virol.*, 78(Pt 1):13-21 (Jan 1997); Furler, S., *et al*, *Gene Ther.*, 8(11):864-873 (June 2001); Klump H., *et al.*, *Gene Ther.*, 8(10):811-817 (May 2001). This 2A peptide is significantly smaller than an IRES, making it well suited for use when space is a limiting factor. However, the selected transgene may encode any biologically active product or other product, e.g., a product desirable for study.

Suitable transgenes may be readily selected by one of skill in the art. The selection of the transgene is not considered to be a limitation of this invention.

2. Regulatory Elements

In addition to the major elements identified above for the minigene, the vector also includes conventional control elements necessary which are operably linked to the transgene in a manner which permits its transcription, translation

and/or expression in a cell transfected with the plasmid vector or infected with the virus produced by the invention. As used herein, “operably linked” sequences include both expression control sequences that are contiguous with the gene of interest and expression control sequences that act in *trans* or at a distance to control the gene of interest.

5 Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation (polyA) signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (i.e., Kozak consensus sequence); sequences that enhance protein stability; and when desired,
10 sequences that enhance secretion of the encoded product. A great number of expression control sequences, including promoters which are native, constitutive, inducible and/or tissue-specific, are known in the art and may be utilized.

Examples of constitutive promoters include, without limitation, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) [see, e.g., Boshart *et al*, *Cell*, **41**:521-530 (1985)], the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1 α promoter [Invitrogen].

Inducible promoters allow regulation of gene expression and can be regulated by exogenously supplied compounds, environmental factors such as temperature, or the presence of a specific physiological state, e.g., acute phase, a particular differentiation state of the cell, or in replicating cells only. Inducible promoters and inducible systems are available from a variety of commercial sources, including, without limitation, Invitrogen, Clontech and Ariad. Many other systems have been described and can be readily selected by one of skill in the art. Examples of inducible promoters regulated by exogenously supplied compounds include, e.g., the zinc-inducible sheep metallothioneine (MT) promoter, the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter, the T7 polymerase promoter system [WO 98/10088]; the ecdysone insect promoter [No *et al*, *Proc. Natl. Acad. Sci. USA*, 93:3346-3351 (1996)], the tetracycline-repressible system [Gossen *et al*, *Proc. Natl. Acad. Sci. USA*, 89:5547-5551 (1992)], the tetracycline-inducible system [Gossen *et al*, *Science*, 268:1766-1769 (1995), see also Harvey *et al*, *Curr. Opin. Chem. Biol.*,

2:512-518 (1998)], the RU486-inducible system [Wang *et al.*, *Nat. Biotech.*, **15**:239-243 (1997) and Wang *et al.*, *Gene Ther.*, **4**:432-441 (1997)] and the rapamycin-inducible system [Magari *et al.*, *J. Clin. Invest.*, **100**:2865-2872 (1997)]. Other types of inducible promoters which may be useful in this context are those which are regulated by a specific physiological state, e.g., temperature, acute phase, a particular differentiation state of the cell, or in replicating cells only.

In another embodiment, the native promoter for the transgene will be used. The native promoter may be preferred when it is desired that expression of the transgene mimic the native expression. The native promoter may be used when expression of the transgene must be regulated temporally or developmentally, or in a tissue-specific manner, or in response to specific transcriptional stimuli. In a further embodiment, other native expression control elements, such as enhancer elements, polyadenylation sites or Kozak consensus sequences may also be used to mimic the native expression.

Another embodiment of the transgene includes a transgene operably linked to a tissue-specific promoter. For instance, if expression in skeletal muscle is desired, a promoter active in muscle should be used. These include the promoters from genes encoding skeletal β -actin, myosin light chain 2A, dystrophin, muscle creatine kinase, as well as synthetic muscle promoters with activities higher than naturally-occurring promoters (see Li *et al.*, *Nat. Biotech.*, **17**:241-245 (1999)). Examples of promoters that are tissue-specific are known for liver (albumin, Miyatake *et al.*, *J. Virol.*, **71**:5124-32 (1997); hepatitis B virus core promoter, Sandig *et al.*, *Gene Ther.*, **3**:1002-9 (1996); alpha-fetoprotein (AFP), Arbuthnot *et al.*, *Hum. Gene Ther.*, **7**:1503-14 (1996)), bone osteocalcin (Stein *et al.*, *Mol. Biol. Rep.*, **24**:185-96 (1997)); bone sialoprotein (Chen *et al.*, *J. Bone Miner. Res.*, **11**:654-64 (1996)), lymphocytes (CD2, Hansal *et al.*, *J. Immunol.*, **161**:1063-8 (1998); immunoglobulin heavy chain; T cell receptor α chain), neuronal such as neuron-specific enolase (NSE) promoter (Andersen *et al.*, *Cell. Mol. Neurobiol.*, **13**:503-15 (1993)), neurofilament light-chain gene (Piccioli *et al.*, *Proc. Natl. Acad. Sci. USA*, **88**:5611-5 (1991)), and the neuron-specific *vgf* gene (Piccioli *et al.*, *Neuron*, **15**:373-84 (1995)), among others.

Optionally, plasmids carrying therapeutically useful transgenes may also include selectable markers or reporter genes may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. Such selectable reporters or marker genes (preferably located outside the viral genome to be rescued by the method of the invention) can be used to signal the presence of the plasmids in bacterial cells, such as ampicillin resistance. Other components of the plasmid may include an origin of replication. Selection of these and other promoters and vector elements are conventional and many such sequences are available [see, e.g., Sambrook et al, and references cited therein].

The combination of the transgene, promoter/enhancer, and 5' and 3' ITRs is referred to as a "minigene" for ease of reference herein. Provided with the teachings of this invention, the design of such a minigene can be made by resort to conventional techniques.

3. Delivery of the Minigene to a Packaging Host Cell

The minigene can be carried on any suitable vector, e.g., a plasmid, which is delivered to a host cell. The plasmids useful in this invention may be engineered such that they are suitable for replication and, optionally, integration in prokaryotic cells, mammalian cells, or both. These plasmids (or other vectors carrying the 5' AAV ITR-heterologous molecule-3'ITR) contain sequences permitting replication of the minigene in eukaryotes and/or prokaryotes and selection markers for these systems. Selectable markers or reporter genes may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. The plasmids may also contain certain selectable reporters or marker genes that can be used to signal the presence of the vector in bacterial cells, such as ampicillin resistance. Other components of the plasmid may include an origin of replication and an amplicon, such as the amplicon system employing the Epstein Barr virus nuclear antigen. This amplicon system, or other similar amplicon components permit high copy episomal replication in the cells. Preferably, the molecule carrying the minigene is transfected into the cell, where it may exist transiently. Alternatively, the minigene (carrying the 5' AAV ITR-heterologous molecule-3' ITR) may be stably integrated into the genome of the host cell, either chromosomally or as an episome. In certain embodiments, the minigene may be

present in multiple copies, optionally in head-to-head, head-to-tail, or tail-to-tail concatamers. Suitable transfection techniques are known and may readily be utilized to deliver the minigene to the host cell.

Generally, when delivering the vector comprising the minigene
5 by transfection, the vector is delivered in an amount from about 5 μ g to about 100 μ g DNA, and preferably about 10 to about 50 μ g DNA to about 1×10^4 cells to about 1×10^{13} cells, and preferably about 10^5 cells. However, the relative amounts of vector DNA to host cells may be adjusted, taking into consideration such factors as the selected vector, the delivery method and the host cells selected.

10 B. *Rep* and *Cap* Sequences

In addition to the minigene, the host cell contains the sequences which drive expression of the AAV9 capsid protein (or a capsid protein comprising a fragment of the AAV9 capsid) in the host cell and rep sequences of the same serotype as the serotype of the AAV ITRs found in the minigene, or a cross-complementing serotype. The AAV *cap* and *rep* sequences may be independently
15 obtained from an AAV source as described above and may be introduced into the host cell in any manner known to one in the art as described above. Additionally, when pseudotyping an AAV vector in an AAV9 capsid, the sequences encoding each of the essential rep proteins may be supplied by AAV9, or the sequences encoding the rep proteins may be supplied by different AAV serotypes (e.g., AAV1, AAV2, AAV3,
20 AAV4, AAV5, AAV6, AAV7). For example, the *rep78/68* sequences may be from AAV2, whereas the *rep52/40* sequences may be from AAV1.

In one embodiment, the host cell stably contains the capsid protein under the control of a suitable promoter, such as those described above.
25 Most desirably, in this embodiment, the capsid protein is expressed under the control of an inducible promoter. In another embodiment, the capsid protein is supplied to the host cell in *trans*. When delivered to the host cell in *trans*, the capsid protein may be delivered via a plasmid which contains the sequences necessary to direct expression of the selected capsid protein in the host cell. Most desirably, when delivered to the host
30 cell in *trans*, the plasmid carrying the capsid protein also carries other sequences required for packaging the rAAV, e.g., the *rep* sequences.

In another embodiment, the host cell stably contains the *rep* sequences under the control of a suitable promoter, such as those described above. Most desirably, in this embodiment, the essential *rep* proteins are expressed under the control of an inducible promoter. In another embodiment, the *rep* proteins are supplied to the host cell in *trans*. When delivered to the host cell in *trans*, the *rep* proteins may be delivered via a plasmid which contains the sequences necessary to direct expression of the selected *rep* proteins in the host cell. Most desirably, when delivered to the host cell in *trans*, the plasmid carrying the capsid protein also carries other sequences required for packaging the rAAV, e.g., the *rep* and *cap* sequences.

Thus, in one embodiment, the *rep* and *cap* sequences may be transfected into the host cell on a single nucleic acid molecule and exist stably in the cell as an episome. In another embodiment, the *rep* and *cap* sequences are stably integrated into the genome of the cell. Another embodiment has the *rep* and *cap* sequences transiently expressed in the host cell. For example, a useful nucleic acid molecule for such transfection comprises, from 5' to 3', a promoter, an optional spacer interposed between the promoter and the start site of the *rep* gene sequence, an AAV *rep* gene sequence, and an AAV *cap* gene sequence.

Optionally, the *rep* and/or *cap* sequences may be supplied on a vector that contains other DNA sequences that are to be introduced into the host cells. For instance, the vector may contain the rAAV construct comprising the minigene. The vector may comprise one or more of the genes encoding the helper functions, e.g., the adenoviral proteins E1, E2a, and E4ORF6, and the gene for VAI RNA.

Preferably, the promoter used in this construct may be any of the constitutive, inducible or native promoters known to one of skill in the art or as discussed above. In one embodiment, an AAV P5 promoter sequence is employed. The selection of the AAV to provide any of these sequences does not limit the invention.

In another preferred embodiment, the promoter for *rep* is an inducible promoter, many of which are discussed above in connection with the transgene regulatory elements. One preferred promoter for *rep* expression is the T7 promoter. The vector comprising the *rep* gene regulated by the T7 promoter and the *cap*

gene, is transfected or transformed into a cell which either constitutively or inducibly expresses the T7 polymerase. See WO 98/10088, published March 12, 1998.

The spacer is an optional element in the design of the vector. The spacer is a DNA sequence interposed between the promoter and the *rep* gene
5 ATG start site. The spacer may have any desired design; that is, it may be a random sequence of nucleotides, or alternatively, it may encode a gene product, such as a marker gene. The spacer may contain genes which typically incorporate start/stop and polyA sites. The spacer may be a non-coding DNA sequence from a prokaryote or eukaryote, a repetitive non-coding sequence, a coding sequence without transcriptional controls or a
10 coding sequence with transcriptional controls. Two exemplary sources of spacer sequences are the λ phage ladder sequences or yeast ladder sequences, which are available commercially, e.g., from Gibco or Invitrogen, among others. The spacer may be of any size sufficient to reduce expression of the *rep78* and *rep68* gene products, leaving the *rep52*, *rep40* and *cap* gene products expressed at normal levels. The length
15 of the spacer may therefore range from about 10 bp to about 10.0 kbp, preferably in the range of about 100 bp to about 8.0 kbp. To reduce the possibility of recombination, the spacer is preferably less than 2 kbp in length; however, the invention is not so limited.

Although the molecule(s) providing *rep* and *cap* may exist in the host cell transiently (i.e., through transfection), it is preferred that one or both
20 of the *rep* and *cap* proteins and the promoter(s) controlling their expression be stably expressed in the host cell, e.g., as an episome or by integration into the chromosome of the host cell. The methods employed for constructing embodiments of this invention are conventional genetic engineering or recombinant engineering techniques such as those described in the references above. While this specification provides illustrative examples
25 of specific constructs, using the information provided herein, one of skill in the art may select and design other suitable constructs, using a choice of spacers, P5 promoters, and other elements, including at least one translational start and stop signal, and the optional addition of polyadenylation sites.

In another embodiment of this invention, the *rep* or *cap* protein
30 may be provided stably by a host cell.

C. The Helper Functions

The packaging host cell also requires helper functions in order to package the rAAV of the invention. Optionally, these functions may be supplied by a herpesvirus. Most desirably, the necessary helper functions are each provided from a human or non-human primate adenovirus source, such as are described herein or and which are available from a variety of sources, including the ATCC. In one currently preferred embodiment, the host cell is provided with and/or contains an E1a gene product, an E1b gene product, an E2a gene product, and/or an E4 ORF6 gene product. The host cell may contain other adenoviral genes such as VAI RNA, but these genes are not required. In a preferred embodiment, no other adenovirus genes or gene functions are present in the host cell.

By "adenoviral DNA which expresses the E1a gene product", it is meant any adenovirus sequence encoding E1a or any functional E1a portion. Adenoviral DNA which expresses the E2a gene product and adenoviral DNA which expresses the E4 ORF6 gene products are defined similarly. Also included are any alleles or other modifications of the adenoviral gene or functional portion thereof. Such modifications may be deliberately introduced by resort to conventional genetic engineering or mutagenic techniques to enhance the adenoviral function in some manner, as well as naturally occurring allelic variants thereof. Such modifications and methods for manipulating DNA to achieve these adenovirus gene functions are known to those of skill in the art.

The adenovirus E1a, E1b, E2a, and/or E4ORF6 gene products, as well as any other desired helper functions, can be provided using any means that allows their expression in a cell. Each of the sequences encoding these products may be on a separate vector, or one or more genes may be on the same vector. The vector may be any vector known in the art or disclosed above, including plasmids, cosmids and viruses. Introduction into the host cell of the vector may be achieved by any means known in the art or as disclosed above, including transfection, infection, electroporation, liposome delivery, membrane fusion techniques, high velocity DNA-coated pellets, viral infection and protoplast fusion, among others. One or more of the adenoviral genes may be stably integrated into the genome of the host cell, stably expressed as episomes, or expressed

transiently. The gene products may all be expressed transiently, on an episome or stably integrated, or some of the gene products may be expressed stably while others are expressed transiently. Furthermore, the promoters for each of the adenoviral genes may be selected independently from a constitutive promoter, an inducible promoter or a native
5 adenoviral promoter. The promoters may be regulated by a specific physiological state of the organism or cell (i.e., by the differentiation state or in replicating or quiescent cells) or by exogenously-added factors, for example.

D. Host Cells And Packaging Cell Lines

The host cell itself may be selected from any biological organism,
10 including prokaryotic (e.g., bacterial) cells, and eukaryotic cells, including, insect cells, yeast cells and mammalian cells. Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549, WEHI, 3T3, 10T1/2, BHK, MDCK, COS 1, COS 7, BSC 1, BSC 40, BMT 10, VERO, WI38, HeLa, 293 cells (which express functional adenoviral E1), Saos, C2C12, L cells,
15 HT1080, HepG2 and primary fibroblast, hepatocyte and myoblast cells derived from mammals including human, monkey, mouse, rat, rabbit, and hamster. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, i.e., fibroblast, hepatocyte, tumor cell, etc. The requirements for the cell used is that it not carry any adenovirus gene other than E1, E2a and/or E4 ORF6;
20 it not contain any other virus gene which could result in homologous recombination of a contaminating virus during the production of rAAV; and it is capable of infection or transfection of DNA and expression of the transfected DNA. In a preferred embodiment, the host cell is one that has *rep* and *cap* stably transfected in the cell.

One host cell useful in the present invention is a host cell stably
25 transformed with the sequences encoding *rep* and *cap*, and which is transfected with the adenovirus E1, E2a, and E4ORF6 DNA and a construct carrying the minigene as described above. Stable *rep* and/or *cap* expressing cell lines, such as B-50 (PCT/US98/19463), or those described in U.S. Patent No. 5,658,785, may also be similarly employed. Another desirable host cell contains the minimum adenoviral DNA
30 which is sufficient to express E4 ORF6. Yet other cell lines can be constructed using the AAV9 *rep* and/or AAV9 *cap* sequences of the invention.

The preparation of a host cell according to this invention involves techniques such as assembly of selected DNA sequences. This assembly may be accomplished utilizing conventional techniques. Such techniques include cDNA and genomic cloning, which are well known and are described in Sambrook et al., cited
5 above, use of overlapping oligonucleotide sequences of the adenovirus and AAV genomes, combined with polymerase chain reaction, synthetic methods, and any other suitable methods which provide the desired nucleotide sequence.

Introduction of the molecules (as plasmids or viruses) into the host cell may also be accomplished using techniques known to the skilled artisan and as
10 discussed throughout the specification. In one embodiment, standard transfection techniques are used, e.g., CaPO₄ transfection or electroporation, and/or infection by hybrid adenovirus/AAV vectors into cell lines such as the human embryonic kidney cell line HEK 293 (a human kidney cell line containing functional adenovirus E1 genes which provides *trans*-acting E1 proteins).

15 The AAV9 based gene therapy vectors which are generated by one of skill in the art are beneficial for gene delivery to selected host cells and gene therapy patients since no neutralization antibodies to AAV9 have been found in the human population. One of skill in the art may readily prepare other rAAV viral vectors containing the AAV9 capsid proteins provided herein using a variety of techniques
20 known to those of skill in the art. One may similarly prepare still other rAAV viral vectors containing AAV9 sequence and AAV capsids of another serotype.

Thus, one of skill in the art will readily understand that the AAV9 sequences of the invention can be readily adapted for use in these and other viral vector systems for *in vitro*, *ex vivo* or *in vivo* gene delivery. Similarly, one of skill in the art can
25 readily select other fragments of the AAV9 genome of the invention for use in a variety of rAAV and non-rAAV vector systems. Such vector systems may include, e.g., lentiviruses, retroviruses, poxviruses, vaccinia viruses, and adenoviral systems, among others. Selection of these vector systems is not a limitation of the present invention.

Thus, the invention further provides vectors generated using the
30 nucleic acid and amino acid sequences of the novel AAV of the invention. Such vectors are useful for a variety of purposes, including for delivery of therapeutic molecules and

for use in vaccine regimens. Particularly desirable for delivery of therapeutic molecules are recombinant AAV containing capsids of the novel AAV of the invention. These, or other vector constructs containing novel AAV sequences of the invention may be used in vaccine regimens, e.g., for co-delivery of a cytokine, or for delivery of the immunogen
5 itself.

V. Recombinant Viruses And Uses Therefor

Using the techniques described herein, one of skill in the art can generate a rAAV having a capsid of a serotype 8 of the invention or having a capsid containing
10 one or more fragments of AAV9. In one embodiment, a full-length capsid from a single serotype, e.g., AAV9 [SEQ ID NO: 2] can be utilized. In another embodiment, a full-length capsid may be generated which contains one or more fragments of AAV9 fused in frame with sequences from another selected AAV serotype, or from heterologous portions of AAV9. For example, a rAAV may contain one or more of the novel
15 hypervariable region sequences of AAV9. Alternatively, the unique AAV9 sequences of the invention may be used in constructs containing other viral or non-viral sequences. Optionally, a recombinant virus may carry AAV9 rep sequences encoding one or more of the AAV9 rep proteins.

A. Delivery of Viruses

20 In another aspect, the present invention provides a method for delivery of a transgene to a host which involves transfecting or infecting a selected host cell with a recombinant viral vector generated with the AAV9 sequences (or functional fragments thereof) of the invention. Methods for delivery are well known to those of skill in the art and are not a limitation of the present invention.

25 In one desirable embodiment, the invention provides a method for AAV9-mediated delivery of a transgene to a host. This method involves transfecting or infecting a selected host cell with a recombinant viral vector containing a selected transgene under the control of sequences which direct expression thereof and AAV9 capsid proteins.

30 Optionally, a sample from the host may be first assayed for the presence of antibodies to a selected AAV serotype. A variety of assay formats for detecting

neutralizing antibodies are well known to those of skill in the art. The selection of such an assay is not a limitation of the present invention. See, e.g., Fisher et al, *Nature Med.*, 3(3):306-312 (March 1997) and W. C. Manning et al, *Human Gene Therapy*, 9:477-485 (March 1, 1998). The results of this assay may be used to determine which AAV vector containing capsid proteins of a particular serotype are preferred for delivery, e.g., by the absence of neutralizing antibodies specific for that capsid serotype.

In one aspect of this method, the delivery of vector with AAV9 capsid proteins may precede or follow delivery of a gene via a vector with a different serotype AAV capsid protein. Thus, gene delivery via rAAV vectors may be used for repeat gene delivery to a selected host cell. Desirably, subsequently administered rAAV vectors carry the same transgene as the first rAAV vector, but the subsequently administered vectors contain capsid proteins of serotypes which differ from the first vector. For example, if a first vector has AAV9 capsid proteins, subsequently administered vectors may have capsid proteins selected from among the other serotypes, including, without limitation, AAV1, AAV2, AAV3, AAV4, AAV8, AAV6, AAV7, and AAV8.

The above-described recombinant vectors may be delivered to host cells according to published methods. The rAAV, preferably suspended in a physiologically compatible carrier, may be administered to a human or non-human mammalian patient. Suitable carriers may be readily selected by one of skill in the art in view of the indication for which the transfer virus is directed. For example, one suitable carrier includes saline, which may be formulated with a variety of buffering solutions (e.g., phosphate buffered saline). Other exemplary carriers include sterile saline, lactose, sucrose, calcium phosphate, gelatin, dextran, agar, pectin, peanut oil, sesame oil, and water. The selection of the carrier is not a limitation of the present invention.

Optionally, the compositions of the invention may contain, in addition to the rAAV and carrier(s), other conventional pharmaceutical ingredients, such as preservatives, or chemical stabilizers. Suitable exemplary preservatives include chlorobutanol, potassium sorbate, sorbic acid, sulfur dioxide, propyl gallate, the parabens, ethyl vanillin, glycerin, phenol, and parachlorophenol. Suitable chemical stabilizers include gelatin and albumin.

The viral vectors are administered in sufficient amounts to transfect the cells and to provide sufficient levels of gene transfer and expression to provide a therapeutic benefit without undue adverse effects, or with medically acceptable physiological effects, which can be determined by those skilled in the medical arts.

5 Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, direct delivery to the liver or lung, orally, intranasally, intratracheally, by inhalation, intravenously, intramuscularly, intraocularly, subcutaneously, intradermally, or by other routes of administration. Routes of administration may be combined, if desired.

10 Dosages of the viral vector will depend primarily on factors such as the condition being treated, the age, weight and health of the patient, and may thus vary among patients. For example, a therapeutically effective human dosage of the viral vector is generally in the range of from about 1 ml to about 100 ml of solution containing concentrations of from about 1×10^9 to 1×10^{16} genomes virus vector. A preferred
15 human dosage may be about 1×10^{13} to 1×10^{16} AAV genomes. The dosage will be adjusted to balance the therapeutic benefit against any side effects and such dosages may vary depending upon the therapeutic application for which the recombinant vector is employed. The levels of expression of the transgene can be monitored to determine the frequency of dosage resulting in viral vectors, preferably AAV vectors containing the
20 minigene. Optionally, dosage regimens similar to those described for therapeutic purposes may be utilized for immunization using the compositions of the invention.

Examples of therapeutic products and immunogenic products for delivery by the AAV9-containing vectors of the invention are provided below. These vectors may be used for a variety of therapeutic or vaccinal regimens, as described herein.
25 Additionally, these vectors may be delivered in combination with one or more other vectors or active ingredients in a desired therapeutic and/or vaccinal regimen.

B. Therapeutic Transgenes

Useful therapeutic products encoded by the transgene include hormones and growth and differentiation factors including, without limitation, insulin, glucagon,
30 growth hormone (GH), parathyroid hormone (PTH), growth hormone releasing factor (GRF), follicle stimulating hormone (FSH), luteinizing hormone (LH), human chorionic

gonadotropin (hCG), vascular endothelial growth factor (VEGF), angiopoietins, angiostatin, granulocyte colony stimulating factor (GCSF), erythropoietin (EPO), connective tissue growth factor (CTGF), basic fibroblast growth factor (bFGF), acidic fibroblast growth factor (aFGF), epidermal growth factor (EGF), platelet-derived growth factor (PDGF), insulin growth factors I and II (IGF-I and IGF-II), any one of the transforming growth factor α superfamily, including TGF α , activins, inhibins, or any of the bone morphogenic proteins (BMP) BMPs 1-15, any one of the heregluin/neuregulin/ARIA/neu differentiation factor (NDF) family of growth factors, nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophins NT-3 and NT-4/5, ciliary neurotrophic factor (CNTF), glial cell line derived neurotrophic factor (GDNF), neurturin, agrin, any one of the family of semaphorins/collapsins, netrin-1 and netrin-2, hepatocyte growth factor (HGF), ephrins, noggin, sonic hedgehog and tyrosine hydroxylase.

Other useful transgene products include proteins that regulate the immune system including, without limitation, cytokines and lymphokines such as thrombopoietin (TPO), interleukins (IL) IL-1 through IL-25 (including, e.g., IL-2, IL-4, IL-6, IL-12 and IL-18], monocyte chemoattractant protein, leukemia inhibitory factor, granulocyte-macrophage colony stimulating factor, Fas ligand, tumor necrosis factors α and β , interferons α , β , and γ , stem cell factor, flk-2/flt3 ligand. Gene products produced by the immune system are also useful in the invention. These include, without limitations, immunoglobulins IgG, IgM, IgA, IgD and IgE, chimeric immunoglobulins, humanized antibodies, single chain antibodies, T cell receptors, chimeric T cell receptors, single chain T cell receptors, class I and class II MHC molecules, as well as engineered immunoglobulins and MHC molecules. Useful gene products also include complement regulatory proteins such as complement regulatory proteins, membrane cofactor protein (MCP), decay accelerating factor (DAF), CR1, CF2 and CD59.

Still other useful gene products include any one of the receptors for the hormones, growth factors, cytokines, lymphokines, regulatory proteins and immune system proteins. The invention encompasses receptors for cholesterol regulation, including the low density lipoprotein (LDL) receptor, high density lipoprotein (HDL) receptor, the very low density lipoprotein (VLDL) receptor, and the scavenger receptor.

The invention also encompasses gene products such as members of the steroid hormone receptor superfamily including glucocorticoid receptors and estrogen receptors, Vitamin D receptors and other nuclear receptors. In addition, useful gene products include transcription factors such as *jun*, *fos*, max, mad, serum response factor (SRF), AP-1, AP2, *myb*, MyoD and myogenin, ETS-box containing proteins, TFE3, E2F, ATF1, ATF2, ATF3, ATF4, ZF5, NFAT, CREB, HNF-4, C/EBP, SP1, CCAAT-box binding proteins, interferon regulation factor (IRF-1), Wilms tumor protein, ETS-binding protein, STAT, GATA-box binding proteins, e.g., GATA-3, and the forkhead family of winged helix proteins.

Other useful gene products include, carbamoyl synthetase I, ornithine transcarbamylase, arginosuccinate synthetase, arginosuccinate lyase, arginase, fumarylacetate hydrolase, phenylalanine hydroxylase, alpha-1 antitrypsin, glucose-6-phosphatase, porphobilinogen deaminase, factor VIII, factor IX, cystathione beta-synthase, branched chain ketoacid decarboxylase, albumin, isovaleryl-coA dehydrogenase, propionyl CoA carboxylase, methyl malonyl CoA mutase, glutaryl CoA dehydrogenase, insulin, beta-glucosidase, pyruvate carboxylate, hepatic phosphorylase, phosphorylase kinase, glycine decarboxylase, H-protein, T-protein, a cystic fibrosis transmembrane regulator (CFTR) sequence, and a dystrophin cDNA sequence. Still other useful gene products include enzymes such as may be useful in enzyme replacement therapy, which is useful in a variety of conditions resulting from deficient activity of enzyme. For example, enzymes that contain mannose-6-phosphate may be utilized in therapies for lysosomal storage diseases (e.g., a suitable gene includes that encoding β -glucuronidase (GUSB)).

Other useful gene products include non-naturally occurring polypeptides, such as chimeric or hybrid polypeptides having a non-naturally occurring amino acid sequence containing insertions, deletions or amino acid substitutions. For example, single-chain engineered immunoglobulins could be useful in certain immunocompromised patients. Other types of non-naturally occurring gene sequences include antisense molecules and catalytic nucleic acids, such as ribozymes, which could be used to reduce overexpression of a target.

Reduction and/or modulation of expression of a gene is particularly desirable for treatment of hyperproliferative conditions characterized by hyperproliferating cells, as are cancers and psoriasis. Target polypeptides include those polypeptides which are produced exclusively or at higher levels in hyperproliferative cells as compared to normal cells. Target antigens include polypeptides encoded by oncogenes such as myb, myc, fyn, and the translocation gene bcr/abl, ras, src, P53, neu, trk and EGRF. In addition to oncogene products as target antigens, target polypeptides for anti-cancer treatments and protective regimens include variable regions of antibodies made by B cell lymphomas and variable regions of T cell receptors of T cell lymphomas which, in some embodiments, are also used as target antigens for autoimmune disease. Other tumor-associated polypeptides can be used as target polypeptides such as polypeptides which are found at higher levels in tumor cells including the polypeptide recognized by monoclonal antibody 17-1A and folate binding polypeptides.

Other suitable therapeutic polypeptides and proteins include those which may be useful for treating individuals suffering from autoimmune diseases and disorders by conferring a broad based protective immune response against targets that are associated with autoimmunity including cell receptors and cells which produce "self"-directed antibodies. T cell mediated autoimmune diseases include Rheumatoid arthritis (RA), multiple sclerosis (MS), Sjögren's syndrome, sarcoidosis, insulin dependent diabetes mellitus (IDDM), autoimmune thyroiditis, reactive arthritis, ankylosing spondylitis, scleroderma, polymyositis, dermatomyositis, psoriasis, vasculitis, Wegener's granulomatosis, Crohn's disease and ulcerative colitis. Each of these diseases is characterized by T cell receptors (TCRs) that bind to endogenous antigens and initiate the inflammatory cascade associated with autoimmune diseases.

C. Immunogenic Transgenes

Alternatively, or in addition, the vectors of the invention may contain AAV9 sequences of the invention and a transgene encoding a peptide, polypeptide or protein which induces an immune response to a selected immunogen. For example, immunogens may be selected from a variety of viral families. Example of desirable viral families against which an immune response would be desirable include, the picornavirus family, which includes the genera rhinoviruses, which are responsible for about 50% of

cases of the common cold; the genera enteroviruses, which include polioviruses, coxsackieviruses, echoviruses, and human enteroviruses such as hepatitis A virus; and the genera aphthoviruses, which are responsible for foot and mouth diseases, primarily in non-human animals. Within the picornavirus family of viruses, target antigens include the VP1, VP2, VP3, VP4, and VPG. Another viral family includes the calcivirus family, which encompasses the Norwalk group of viruses, which are an important causative agent of epidemic gastroenteritis. Still another viral family desirable for use in targeting antigens for inducing immune responses in humans and non-human animals is the togavirus family, which includes the genera alphavirus, which include Sindbis viruses, Ross River virus, and Venezuelan, Eastern & Western Equine encephalitis, and rubivirus, including Rubella virus. The flaviviridae family includes dengue, yellow fever, Japanese encephalitis, St. Louis encephalitis and tick borne encephalitis viruses. Other target antigens may be generated from the Hepatitis C or the coronavirus family, which includes a number of non-human viruses such as infectious bronchitis virus (poultry), porcine transmissible gastroenteric virus (pig), porcine hemagglutinating encephalomyelitis virus (pig), feline infectious peritonitis virus (cats), feline enteric coronavirus (cat), canine coronavirus (dog), and human respiratory coronaviruses, which may cause the common cold and/or non-A, B or C hepatitis. Within the coronavirus family, target antigens include the E1 (also called M or matrix protein), E2 (also called S or Spike protein), E3 (also called HE or hemagglutinin-esterase) glycoprotein (not present in all coronaviruses), or N (nucleocapsid). Still other antigens may be targeted against the rhabdovirus family, which includes the genera vesiculovirus (e.g., Vesicular Stomatitis Virus), and the general lyssavirus (e.g., rabies). Within the rhabdovirus family, suitable antigens may be derived from the G protein or the N protein. The family filoviridae, which includes hemorrhagic fever viruses such as Marburg and Ebola virus may be a suitable source of antigens. The paramyxovirus family includes parainfluenza Virus Type 1, parainfluenza Virus Type 3, bovine parainfluenza Virus Type 3, rubulavirus (mumps virus, parainfluenza Virus Type 2, parainfluenza virus Type 4, Newcastle disease virus (chickens), rinderpest, morbillivirus, which includes measles and canine distemper, and pneumovirus, which includes respiratory syncytial virus. The influenza virus is classified within the family orthomyxovirus and is a suitable source of

antigen (e.g., the HA protein, the NI protein). The bunyavirus family includes the genera bunyavirus (California encephalitis, La Crosse), phlebovirus (Rift Valley Fever), hantavirus (puremala is a hemahagin fever virus), nairovirus (Nairobi sheep disease) and various unassigned bungaviruses. The arenavirus family provides a source of antigens
5 against LCM and Lassa fever virus. The reovirus family includes the genera reovirus, rotavirus (which causes acute gastroenteritis in children), orbiviruses, and cultivirus (Colorado Tick fever, Lebombo (humans), equine encephalosis, blue tongue). The retrovirus family includes the sub-family oncorivirinal which encompasses such human and veterinary diseases as feline leukemia virus, HTLVI and HTLVII, lentivirinal (which
10 includes HIV, simian immunodeficiency virus, feline immunodeficiency virus, equine infectious anemia virus, and spumavirinal). The papovavirus family includes the sub-family polyomaviruses (BKU and JCU viruses) and the sub-family papillomavirus (associated with cancers or malignant progression of papilloma). The adenovirus family includes viruses (EX, AD7, ARD, O.B.) which cause respiratory disease and/or enteritis.
15 The parvovirus family feline parvovirus (feline enteritis), feline panleucopeniavirus, canine parvovirus, and porcine parvovirus. The herpesvirus family includes the sub-family alphaherpesvirinae, which encompasses the genera simplexvirus (HSVI, HSVII), varicellovirus (pseudorabies, varicella zoster) and the sub-family betaherpesvirinae, which includes the genera cytomegalovirus (HCMV,
20 muromegalovirus) and the sub-family gammaherpesvirinae, which includes the genera lymphocryptovirus, EBV (Burkitts lymphoma), infectious rhinotracheitis, Marek's disease virus, and rhadinovirus. The poxvirus family includes the sub-family chordopoxvirinae, which encompasses the genera orthopoxvirus (Variola major (Smallpox) and Vaccinia (Cowpox)), parapoxvirus, avipoxvirus, capripoxvirus,
25 leporipoxvirus, suipoxvirus, and the sub-family entomopoxvirinae. The hepadnavirus family includes the Hepatitis B virus. One unclassified virus which may be suitable source of antigens is the Hepatitis delta virus. Another virus which is a source of antigens is Nippan Virus. Still other viral sources may include avian infectious bursal disease virus and porcine respiratory and reproductive syndrome virus. The alphavirus
30 family includes equine arteritis virus and various Encephalitis viruses.

The present invention may also encompass immunogens which are useful to immunize a human or non-human animal against other pathogens including bacteria, fungi, parasitic microorganisms or multicellular parasites which infect human and non-human vertebrates, or from a cancer cell or tumor cell. Examples of bacterial pathogens include pathogenic gram-positive cocci include pneumococci; staphylococci (and the toxins produced thereby, e.g., enterotoxin B); and streptococci. Pathogenic gram-negative cocci include meningococcus; gonococcus. Pathogenic enteric gram-negative bacilli include enterobacteriaceae; pseudomonas, acinetobacteria and eikenella; melioidosis; salmonella; shigella; haemophilus; moraxella; *H. ducreyi* (which causes chancroid); brucella species (brucellosis); *Francisella tularensis* (which causes tularemia); *Yersinia pestis* (plague) and other yersinia (pasteurella); streptobacillus moniliformis and spirillum; Gram-positive bacilli include listeria monocytogenes; erysipelothrix rhusiopathiae; *Corynebacterium diphtheria* (diphtheria); cholera; *B. anthracis* (anthrax); donovanosis (granuloma inguinale); and bartonellosis. Diseases caused by pathogenic anaerobic bacteria include tetanus; botulism (*Clostridium botulinum* and its toxin); *Clostridium perfringens* and its epsilon toxin; other clostridia; tuberculosis; leprosy; and other mycobacteria. Pathogenic spirochetal diseases include syphilis; treponematoses: yaws, pinta and endemic syphilis; and leptospirosis. Other infections caused by higher pathogen bacteria and pathogenic fungi include glanders (*Burkholderia mallei*); actinomycosis; nocardiosis; cryptococcosis, blastomycosis, histoplasmosis and coccidioidomycosis; candidiasis, aspergillosis, and mucormycosis; sporotrichosis; paracoccidioidomycosis, petriellidiosis, torulopsosis, mycetoma and chromomycosis; and dermatophytosis. Rickettsial infections include Typhus fever, Rocky Mountain spotted fever, Q fever (*Coxiella burnetti*), and Rickettsialpox. Examples of mycoplasma and chlamydial infections include: mycoplasma pneumoniae; lymphogranuloma venereum; psittacosis; and perinatal chlamydial infections. Pathogenic eukaryotes encompass pathogenic protozoans and helminths and infections produced thereby include: amebiasis; malaria; leishmaniasis; trypanosomiasis; toxoplasmosis; *Pneumocystis carinii*; *Trichans*; *Toxoplasma gondii*; babesiosis; giardiasis; trichinosis; filariasis; schistosomiasis; nematodes; trematodes or flukes; and cestode (tapeworm) infections.

Many of these organisms and/or the toxins produced thereby have been identified by the Centers for Disease Control [(CDC), Department of Health and Human Services, USA], as agents which have potential for use in biological attacks. For example, some of these biological agents, include, *Bacillus anthracis* (anthrax),
5 *Clostridium botulinum* and its toxin (botulism), *Yersinia pestis* (plague), variola major (smallpox), *Francisella tularensis* (tularemia), and viral hemorrhagic fevers [filoviruses (e.g., Ebola, Marburg], and arenaviruses [e.g., Lassa, Machupo]], all of which are currently classified as Category A agents; *Coxiella burnetti* (Q fever); *Brucella* species (brucellosis), *Burkholderia mallei* (glanders), *Burkholderia pseudomallei* (meloidosis),
10 *Ricinus communis* and its toxin (ricin toxin), *Clostridium perfringens* and its toxin (epsilon toxin), *Staphylococcus* species and their toxins (enterotoxin B), *Chlamydia psittaci* (psittacosis), water safety threats (e.g., *Vibrio cholerae*, *Cryptosporidium parvum*), Typhus fever (*Richettsia powazekii*), and viral encephalitis (alphaviruses, e.g., Venezuelan equine encephalitis; eastern equine encephalitis; western equine
15 encephalitis); all of which are currently classified as Category B agents; and Nipah virus and hantaviruses, which are currently classified as Category C agents. In addition, other organisms, which are so classified or differently classified, may be identified and/or used for such a purpose in the future. It will be readily understood that the viral vectors and other constructs described herein are useful to deliver antigens from these organisms,
20 viruses, their toxins or other by-products, which will prevent and/or treat infection or other adverse reactions with these biological agents.

Administration of the vectors of the invention to deliver immunogens against the variable region of the T cells elicit an immune response including CTLs to eliminate those T cells. In rheumatoid arthritis (RA), several specific variable regions of
25 T cell receptors (TCRs) which are involved in the disease have been characterized. These TCRs include V-3, V-14, V-17 and V-17. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in RA. In multiple sclerosis (MS), several specific variable regions of TCRs which are involved in the disease have been characterized.
30 These TCRs include V-7 and V-10. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target

T cells involved in MS. In scleroderma, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-6, V-8, V-14 and V-16, V-3C, V-7, V-14, V-15, V-16, V-28 and V-12. Thus, delivery of a nucleic acid molecule that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in scleroderma.

Thus, a rAAV9-derived recombinant viral vector of the invention provides an efficient gene transfer vehicle which can deliver a selected transgene to a selected host cell *in vivo* or *ex vivo* even where the organism has neutralizing antibodies to one or more AAV serotypes. In one embodiment, the rAAV and the cells are mixed *ex vivo*; the infected cells are cultured using conventional methodologies; and the transduced cells are re-infused into the patient.

The vectors of the invention are particularly well suited to gene delivery for therapeutic purposes and for immunization, including inducing protective immunity. Further, the vectors of the invention may also be used for production of a desired gene product *in vitro*. For *in vitro* production, a desired product (e.g., a protein) may be obtained from a desired culture following transfection of host cells with a rAAV containing the molecule encoding the desired product and culturing the cell culture under conditions which permit expression. The expressed product may then be purified and isolated, as desired. Suitable techniques for transfection, cell culturing, purification, and isolation are known to those of skill in the art.

The following examples illustrate several aspects and embodiments of the invention.

EXAMPLES

Example 1: Production of Recombinant AAV9 Viral Genomes Equipped With AAV2 ITRs

Chimeric packaging constructs are generated by fusing AAV2 rep with cap sequences of novel AAV serotypes. These chimeric packaging constructs are used, initially, for pseudotyping recombinant AAV genomes carrying AAV2 ITRs by triple

transfection in 293 cell using Ad5 helper plasmid. These pseudotyped vectors are used to evaluate performance in transduction-based serological studies and evaluate gene transfer efficiency of novel AAV serotypes in different animal models including NHP and rodents, before intact and infectious viruses of these novel serotypes are isolated.

5 A. *pAAV2GFP*

The AAV2 plasmid which contains the AAV2 ITRs and green fluorescent protein expressed under the control of a constitutive promoter. This plasmid contains the following elements: the AAV2 ITRs, a CMV promoter, the GFP coding sequences,

10 B. *Cloning of trans plasmid*

To construct the chimeric trans-plasmid for production of recombinant pseudotyped AAV9 vectors, p5E18 plasmid (Xiao *et al.*, 1999, *J. Virol* 73:3994-4003) was partially digested with Xho I to linearize the plasmid at the Xho I site at the position of 3169 bp only. The Xho I cut ends were then filled in and ligated back.

15 This modified p5E18 plasmid was restricted with Xba I and Xho I in a complete digestion to remove the AAV2 cap gene sequence and replaced with a 2267 bp Spe I/Xho I fragment containing the AAV9 cap gene which was isolated from pCRAAV9.6-5+15-4 plasmid.

The resulting plasmid contains the AAV2 rep sequences for Rep78/68 under the control of the AAV2 P5 promoter, and the AAV2 rep sequences for Rep52/40 under the control of the AAV2 P19 promoter. The AAV9 capsid sequences are under the control of the AAV2 P40 promoter, which is located within the Rep sequences. This plasmid further contains a spacer 5' of the rep ORF.

Alternatively, a similar plasmid can be constructed which utilizes the AAV9 rep sequences and the native AAV9 promoter sequences. This plasmid is then used for production of rAAV9, as described herein.

C. *Production of Pseudotyped rAAV*

The rAAV particles (AAV2 vector in AAV9 capsid) are generated using an adenovirus-free method. Briefly, the cis plasmid (pAAV2.1 lacZ plasmid containing AAV2 ITRs), and the trans plasmid pCRAAV9 6-5+15-4 (containing the AAV2 rep and

AAV9cap) and a helper plasmid, respectively, are simultaneously co-transfected into 293 cells in a ratio of 1:1:2 by calcium phosphate precipitation.

For the construction of the pAd helper plasmids, pBG10 plasmid was purchased from Microbix (Canada). A RsrII fragment containing L2 and L3 is deleted from pBHG10, resulting in the first helper plasmid, pAd F13. Plasmid Ad F1 was constructed by cloning Asp700/SalI fragment with a PmeI/SgfI deletion, isolating from pBHG10, into Bluescript. MLP, L2, L2 and L3 were deleted in the pAd F1. Further deletions of a 2.3 kb NruI fragment and, subsequently, a 0.5 kb RsrII/NruI fragment generated helper plasmids pAd F5 and pAd F6, respectively. The helper plasmid, termed p F6, provides the essential helper functions of E2a and E4 ORF6 not provided by the E1-expressing helper cell, but is deleted of adenoviral capsid proteins and functional E1 regions).

Typically, 50 µg of DNA (cis:trans:helper) is transfected onto a 150 mm tissue culture dish. The 293 cells are harvested 72 hours post-transfection, sonicated and treated with 0.5% sodium deoxycholate (37 C for 10 min.) Cell lysates are then subjected to two rounds of a CsCl gradient. Peak fractions containing rAAV vector are collected, pooled and dialyzed against PBS.

Example 2 – Mouse Model of Familial Hypercholesterolemia

The following experiment will demonstrate that the AAV2/9 rAAV constructed as described herein delivers the LDL receptor and expresses LDL receptor in an amount sufficient to reduce the levels of plasma cholesterol and triglycerides in animal models of familial hypercholesterolemia.

The rAAV2/9 constructs used in the following experiment contain the cDNA for murine VLDL receptor (LDLR) driven from a β-actin promoter and enhanced from the cytomegalovirus enhancer (AAV-CB-VLDLR). LDL receptor-deficient mice (Jackson Laboratories) on a high cholesterol diet which exhibit the symptoms of familial hypercholesterolemia, (FH mice) are infused intravenously with 4×10^{12} AAV-CB-LDLR or a saline control and the plasma cholesterol level was monitored.

The plasma cholesterol and triglyceride profiles of FH mice infused with VLDLR virus are determined as follows. The mice treated as described above are fasted for six (6) hours, and blood samples collected by retro-orbital venous plexus puncture with heparinized capillary tubes. Plasma is separated by centrifugation. The concentration of total cholesterol and triglyceride is determined using kits purchased from Wako Chemicals.

Mice in the control group show gradually increased cholesterol levels until reaching a steady state in about 4 weeks post injection. In mice receiving a LDLR virus, plasma cholesterol levels steadily decrease starting 3 weeks post-virus administration.

A cholesterol profile of plasma from the FH mice infused with LDLR virus is obtained by FPLC analysis. Mouse plasma is pooled from 4 mice in each group. Plasma samples are isolated from control and AAV-VLDLR infused animals at pre-injection (PI) or 2 months (2M) post virus infusion and are analyzed by FPLC fractionation followed by cholesterol assay.

Example 3 - *In vivo* Transduction with AAV9 Serotype Vectors

The performance of vector based on the new AAV9 serotype is evaluated in murine models of muscle and liver-directed gene transfer and compared to vectors based on the known serotypes AAV1, AAV2 and AAV5. Vectors expressing secreted proteins are used to quantitate relative transduction efficiencies between different serotypes through ELISA analysis of sera. The cellular distribution of transduction within the target organ is evaluated using lacZ expressing vectors and X-gal histochemistry.

For this experiment, recombinant AAV genomes, AAV2CBhA1AT, AAV2AlbA1AT, AAV2CMVrhCG, AAV2TBGrhCG, AAV2TBGcFIX, AAV2CMVLacZ and AAV2TBGLacZ are packaged with AAV9, or AAV1, AAV2, or AAV5 capsid proteins. In all constructs, minigene cassettes are flanked with AAV2 ITRs. cDNAs of human α -antitrypsin (A1AT) [Xiao, W., *et al.*, (1999) *J Virol* 73, 3994-4003] β -subunit of rhesus monkey choriogonadotropic hormone (CG) [Zoltick, P. W. &

Wilson, J. M. (2000) *Mol Ther* 2, 657-9] canine factor IX [Wang, L., *et al.*, (1997) *Proc Natl Acad Sci U S A* 94, 11563-6] and bacterial β -galactosidase (i.e., Lac Z) genes are used as reporter genes. For liver-directed gene transfer, either mouse albumin gene promoter (Alb) [Xiao, W. (1999), cited above] or human thyroid hormone binding globulin gene promoter (TBG) [Wang (1997), cited above] is used to drive liver specific expression of reporter genes. In muscle-directed gene transfer experiments, either cytomegalovirus early promoter (CMV) or chicken β -actin promoter with CMV enhancer (CB) is employed to direct expression of reporters.

For muscle-directed gene transfer, vectors are injected into the right tibialis anterior of 4-6 week old NCR nude or C57BL/6 mice (Taconic, Germantown, NY). In liver-directed gene transfer studies, vectors are infused intraportally into 7-9 week old NCR nude or C57BL/6 mice (Taconic, Germantown, NY). Serum samples are collected intraorbitally at different time points after vector administration. Muscle and liver tissues are harvested at different time points for cryosectioning and Xgal histochemical staining from animals that received the lacZ vectors. For the re-administration experiment, C56BL/6 mice initially received AAV2/1, 2/2, 2/5, and 2/9 CBA1AT vectors intramuscularly and followed for A1AT gene expression for 7 weeks. Previous studies indicated that immune competent C57BL/6 mice elicit limited humoral responses to the human A1AT protein when expressed from AAV vectors [Xiao, W., *et al.*, (1999) *J Virol* 73, 3994-4003]. Animals are then treated with AAV2/9 TBGcFIX intraportally and studied for cFIX gene expression.

ELISA based assays are performed to quantify serum levels of hA1AT, rhCG and cFIX proteins as described previously [Gao, G. P., *et al.*, (1996) *J Virol* 70, 8934-43; Zoltick, P. W. & Wilson, J. M. (2000) *Mol Ther* 2, 657-9; Wang, L., *et al.*, *Proc Natl Acad Sci U S A* 94, 11563-6]. The experiments are completed when animals are sacrificed for harvest of muscle and liver tissues for DNA extraction and quantitative analysis of genome copies of vectors present in target tissues by TaqMan using the same set of primers and probe as in titration of vector preparations [Zhang, Y., *et al.*, (2001) *Mol Ther* 3, 697-707].

All publications cited in this specification are incorporated herein by reference. While the invention has been described with reference to particularly preferred embodiments, it will be appreciated that modifications can be made without departing from the spirit of the invention. Such modifications are intended to fall within the scope
5 of the claims.

What is claimed is:

1. An isolated adeno-associated virus (AAV) comprising an AAV9 capsid having an amino acid sequence of SEQ ID NO:2.
2. The isolated AAV according to claim 1, wherein said virus comprises the nucleic acid sequence of SEQ ID NO:1.
3. The isolated AAV according to claim 1, wherein said AAV further comprises a minigene having AAV inverted terminal repeats and a heterologous gene operably linked to regulatory sequences which direct its expression in a host cell.
4. A protein comprising an AAV9 protein or a fragment thereof selected from the group consisting of:
 - (a) an AAV9 capsid protein or fragment thereof, selected from the group consisting of:
 - vp1 capsid protein, amino acids (aa) 1 to 736;
 - vp2 capsid protein, aa 138 to 736;
 - vp3 capsid protein, aa 203 to 736;
 - a fragment encompassing hypervariable region (HVR)1 through 12 or a smaller fragment thereof selected the group consisting of: aa 146 to 152; aa 182 to 187; aa 262 to 264; aa 263 to 266; aa 263 to 266; aa 381 to 383; 383 to 385; aa 450 to 474; aa 451 to 475; aa 490 to 495; aa 491 to 496; aa500 to 504; aa 501 to 505; aa 514 to 522; aa 533 to 554; aa 534 to 555; aa 581 to 594; aa 583 to 596; aa 658 to 667; aa 660 to 669; and aa 705 to 719; aa 707 to 723;
 - aa 24 to 42, aa 25 to 28; aa 81 to 85; aa133 to 165; aa 134 to 165;
 - aa 137 to 143; aa 154 to 156; aa 194 to 208; aa 261 to 274; aa 262 to 274; aa 171 to 173;
 - aa 185 to 198; aa 413 to 417; aa 449 to 478; aa 494 to 525; aa 534 to 571; aa 581 to 601;
 - aa 660 to 671; aa 709 to 723; and

aa 1 to 184, aa 199 to 259; aa 274 to 446; aa 603 to 659; aa 670 to 706; aa 724 to 736; aa 185 to 198; aa 260 to 273; aa447 to 477; aa495 to 602; aa603 to 659; aa 660 to 669; and aa707 to 723,

wherein the amino acid numbers are those of the capsid of AAV9,
SEQ ID NO:2;

and

(b) an AAV9 rep protein or fragment thereof selected from the group consisting of:

aa 1 to 623; aa 1 to 102; aa 103 to 140; aa 141 to 173; aa 174 to 226; aa 227 to 275; aa 276 to 374; aa 375 to 383; aa 384 to 446; aa 447 to 542; aa 543 to 555; and aa 556 to 623, of SEQ ID NO: 3.

5. An artificial adeno-associated virus (AAV) capsid protein comprising one or more of the AAV9 capsid protein fragments according to claim 4a.

6. A recombinant adeno-associated virus (AAV) comprising an artificial capsid according to claim 5.

7. A molecule comprising a nucleic acid sequence encoding a protein according to claim 4.

8. The molecule according to claim 7, wherein said nucleic acid sequence is selected from the group consisting of:

vp1, nt 2121 to 4323;
vp2, nt 2532 to 4323; and
vp 3, nt 2730 to 4323,

wherein the nucleotides numbers are of AAV9, SEQ ID NO:1.

9. The molecule according to claim 7 or claim 8, wherein said molecule comprises an AAV sequence encoding an AAV capsid protein and a functional AAV rep protein.

10. The molecule according to any of claims 7 to 9, wherein said nucleic acid sequence comprises a sequence selected from the group consisting of: nucleic acids 900 to 2099 of SEQ ID NO:1, nucleic acids 2227 to 2099 of SEQ ID NO:1; nucleic acids 905 to 2104 of SEQ ID NO:1; and nucleic acids 237 to 2104 of SEQ ID NO:1.

11. The molecule according to claim 9, wherein said molecule comprises a cap protein or a functional AAV rep gene from a serotype selected from the group consisting of AAV1, AAV2, AAV3, AAV4, AAV5 and AAV6.

12. The molecule according to any of claims 7 to 11, wherein said molecule is a plasmid.

13. A method of generating a recombinant adeno-associated virus (AAV) comprising an AAV serotype capsid comprising the steps of culturing a host cell containing: (a) a molecule encoding an AAV capsid protein; (b) a functional rep gene; (c) a minigene comprising AAV inverted terminal repeats (ITRs) and a transgene; and (d) sufficient helper functions to permit packaging of the minigene into the AAV capsid protein, wherein said host cell comprises a molecule according to any of claims 7 to 12.

14. A host cell transfected with an adeno-associated virus according to any of claims 1 to 3 or claim 6 or a molecule according to any of claims 7 to 12.

15. A composition comprising an AAV according to any of claims 1 to 3 or claim 6, and a physiologically compatible carrier.

16. A method of delivering a transgene to a cell, said method comprising the step of contacting the cell with an AAV according to any of claims 1 to 3 or claim 6, wherein said rAAV comprises the transgene.

17. Use of an adeno-associated virus according to any of claims 1 to 3 or claim 6 or a molecule according to any of claims 7 to 11 in preparing a medicament for delivery of a transgene to a cell.

Fig. 1A

cagagagggga gtggccaact ccatcactag gggtaatcgc gaagcgcctc ccacgctgcc 60
 gcgtcagcgc tgacgtagat tacgtcatag gggagtggtc ctgtattagc tgtcacgtga 120
 gtgcttttgc gacattttgc gacaccacat ggccatttga ggtatatatg gccgagtga 180
 cgagcaggat ctccattttg accgcgaaat ttgaacgagc agcagccatg ^{Rep68/78 start} ccgggcttct 240
 acgagattgt gatcaagggtg ccgagcgacc tggacgagca cctgccgggc atttctgact 300
 cttttgtgaa ctgggtggcc gagaaggaat gggagctgcc cccggattct gacatggatc 360
 ggaatctgat cgagcaggca cccctgaccg tggccgagaa gctgcagcgc gacttcctgg 420
 tccaatggcg ccgctgagt aaggcccgagg aggcctctt ctttgttcag ttcgagaagg 480
 gcgagagcta ctttcacctg cacgttctgg tcgagaccac gggggcaag tccatgggtgc 540
 taggccgctt cctgagtcag attcgggaga agctgggtcca gaccatctac cgcgggatcg 600
 agccgaccct gcccaactgg ttcgcggtga ccaagacgcg taatggcgcc ggcgggggga 660
 acaagggtggg ggacgagtgc tacatcccca actacctct gcccaagact cagcccgagc 720
 tgcagtgggc gtggactaac atggaggagt atataagcgc gtgcttgaac ctggccgagc 780
 gcaaacggct cgtggcgag cacctgacct acgtcagcca gacgcaggag cagaacaagg 840
 agaatctgaa ccccaattct gacgcgcccg tgatcaggtc aaaaacctcc gcgcgctaca 900
^{Rep40/52 start}
 tgagctggg cggtgggtg gtggaccggg gcatcacctc cgagaagcag tggatccagg 960
 aggaccaggc ctgtacatc tccttcaacg ccgcctcaa ctgcggtcc cagatcaagg 1020
 ccgcgctgga caatgccggc aagatcatgg cgctgaccaa atccgcgcc gactacctgg 1080
 taggcccttc acttcgggtg gacattacgc agaaccgat ctaccgcatc ctgcagctca 1140
 acggctacga ccctgcctac gccggctccg tctttctcgg ctgggcacaa aagaagtgc 1200
 ggaaacgcaa caccatctgg ctgtttgggc cggccaccac gggaaagacc aacatcgag 1260
 aagccattgc ccacgccgtg cccttctacg gctgcgtcaa ctggaccaat gagaactttc 1320
 ccttcaacga ttgcgtcgac aagatgggtga tctggtggga ggagggaag atgacggcca 1380
 aggtcgtgga gtccgccaag gccattctcg gcggcagcaa ggtgcgctg gacaaaaagt 1440
 gcaagtcgtc cgccagatc gacccactc ccgtgatcgt cacctccaac accaacaagt 1500

Fig. 1B

gcgccgtgat tgacgggaac agcaccacct tcgagcacca gcagcctctc caggaccgga 1560
 tgtttaagtt cgaactcacc cgccgtctgg agcacgactt tggcaagggtg acaaagcagg 1620
 aagtcaaaga gttcttccgc tgggccagtg atcacgtgac cgaggtggcg catgagtttt 1680
 acgtcagaaa gggcgaggcc agcaaaagac ccgccccga tgacgcggat aaaagcgagc 1740
 ccaagcgggc ctgccctca gtgcgggatc catcgacgtc agacgcggaa ggagctccgg 1800
 tggactttgc cgacaggtac caaaacaaat gttctcgtca cgcgggcatg cttcagatgc 1860
 tgcttccttg caaaacgtgc gagagaatga atcagaattt caacatttgc ttcacacacg 1920
 gggtcagaga ctgctcagag tgtttcccg gcgtgtcaga atctcaaccg gtcgtcagaa 1980
 agaggacgta tcggaaactc tgtgcgattc atcatctgct ggggcgggct cccgagattg 2040
 cttgctcggc ctgcgatctg gtcaacgtgg acctggatga ctgtgtttct ^{Rep 78 stop} gagcaat^{aaa} 2100
 tgacttaa^{ac} ^{vp1 start} cagg^{tat}ggc tgccgatggg tatcttcag attggctcga ggacaacctc 2160
 tctgagggca ttcgcgagtg gtgggaacctg aaacctggag ccccgaaacc caaagccaac 2220
 cagcaaaagc aggacgacgg ccggggtctg gtgcttcctg gctacaagta cctcggaacc 2280
 ttcaacggac tcgacaaggg ggagcccgtc aacgcggcgg acgcagcggc cctcgagcac 2340
 gacaaggcct acgaccagca gctcaaagcg ggtgacaatc cgtacctgcg gtataaccac 2400
 gccgacgccg agtttcagga gcgtctgcaa gaagatacgt cttttggggg caacctcggg 2460
 cgagcagtct tcaggccaa gaagcgggtt ctcgaacctc tcgggtctggg tgaggaaggc 2520
 gctaag^{ac}gg ^{vp2 start} ctcctgaaa gaagagaccg gtagagcagt cacc^{ca}aga accagactca 2580
 tcctcgggca tcggcaa^{at}c aggccagcag cccgctaaaa agagactcaa ttttggtcag 2640
 actggcgact cagagtcagt cccgaccca caacctctcg gagaacctc agaagcccc 2700
 tcagg^tctgg ^{vp3 start} gacctaatac ^{aat}ggcttca ggcggtggcg ctccaatggc agacaataac 2760
 gaaggcggc acggagtggg taattcctcg ggaaattggc attgcgattc cacatggctg 2820
 ggggacagag tcatcaccac cagcaccgga acctgggcat tgccaccta caacaaccac 2880
 ctctacaagc aaatctccaa tggaacatcg ggaggaagca ccaacgacaa cacctacttt 2940

Fig. 1C

```

ggctacagca cccctgggg gtattttgac ttcaacagat tccactgcca cttctcacca 3000
cgtgactggc agcgactcat caacaacaac tggggattcc ggccaaagag actcaacttc 3060
aagctgttca acatccaggt caaggagggt acgacgaacg aaggcaccaa gaccatcgcc 3120
aataacctta ccagcacctg ccagggtctt acggactcgg agtaccagct accgtacgtc 3180
ctaggctctg cccaccaagg atgcctgcca cgttttctg cagacgtctt catggttctt 3240
cagtacggct acctgacgct caacaatgga agtcaagcgt taggacgttc ttctttctac 3300
tgtctggaat acttcccttc tcagatgctg agaaccggca acaactttca gttcagctac 3360
actttcgagg acgtgccttt ccacagcagc tacgcacaca gccagagtct agatcgactg 3420
atgaaccccc tcatcgacca gtacctatac tacctggta gaacacagac aactggaact 3480
gggggaactc aaactttggc attcagccaa gcaggcccta gctcaatggc caatcaggct 3540
agaaactggg taccogggcc ttgctaccgt cagcagcgcg tctccacaac caccaaccaa 3600
aataacaaca gcaactttgc gtggacggga gctgctaaat tcaagctgaa cgggagagac 3660
tcgctaata ga atcctggcgt ggctatggca tcgcacaaag acgacgagga ccgcttcttt 3720
ccatcaagtg gcgttctcat atttggaag caaggagccg ggaacgatgg agtcgactac 3780
agccagggtc tgattacaga tgaggaagaa attaaagcca ccaaccctgt agccacagag 3840
gaatacggag cagtggccat caacaaccag gccgctaaca cgcaggcgca aactggactt 3900
gtgcataacc agggagtatt tcttggtatg gtctggcaga accgggacgt gtacctgcag 3960
ggcctatatt gggtataaat acctcacaca gatggcaact ttcacccgtc tctctgatg 4020
ggtggatttg gactgaaaca cccacctcca cagattctaa ttaaaaatac accagtgccg 4080
gcagatcctc ctcttacctt caatcaagcc aagctgaact ctttcacac gcagtacagc 4140
acgggacaag tcagcgtgga aatcgagtgg gagctgcaga aagaaaacag caagcgtgg 4200
aatccagaga tccagtatac ttcaaactac taaaaatcta caaatgtgga ctttgctgtc 4260
aataccgaag gtgtttactc tgagcctcgc ccattggta ctggttacct caccgtaat 4320
    vpl-3 stop      polyA
    ttgtaattgc ctgttaatc ataaacccggt taattcggtt cagttgaact ttggtctctg 4380
cg                                                                 4382

```

Fig. 2A

	1				50
AAV_2	MAADGYLPDW	LEDTLSEGIR	QWWKLKPGPP	PPKPAERHKD	DSRGLVLPGY
AAV_8	MAADGYLPDW	LEDNLSEGIR	EWWALKPGAP	KPKANQQKQD	DGRGLVLPGY
AAV_1	MAADGYLPDW	LEDNLSEGIR	EWDDLKPGAP	KPKANQQKQD	DGRGLVLPGY
AAV_3	MAADGYLPDW	LEDNLSEGIR	EWWALKPGVP	QPKANQQHQD	NRRGLVLPGY
AAV_9	MAADGYLPDW	LEDNLSEGIR	EWDDLKPGAP	KPKANQQKQD	DGRGLVLPGY
	51				100
AAV_2	KYLGPFNGLD	KGEPVNEADA	AALEHDKAYD	RQLDSGDNPY	LKYNHADAEF
AAV_8	KYLGPFNGLD	KGEPVNAADA	AALEHDKAYD	QQLQAGDNPY	LRYNHADAEF
AAV_1	KYLGPFNGLD	KGEPVNAADA	AALEHDKAYD	QQLKAGDNPY	LRYNHADAEF
AAV_3	KYLGPFNGLD	KGEPVNEADA	AALEHDKAYD	QQLKAGDNPY	LKYNHADAEF
AAV_9	KYLGPFNGLD	KGEPVNAADA	AALEHDKAYD	QQLKAGDNPY	LRYNHADAEF
	101				150
AAV_2	QERLKEDTSF	GGNLGRAVFQ	AKKRVLEPLG	LVEEPVKTAP	GKKRPVEHSP
AAV_8	QERLQEDTSF	GGNLGRAVFQ	AKKRVLEPLG	LVEEGAKTAP	GKKRPVEPSP
AAV_1	QERLQEDTSF	GGNLGRAVFQ	AKKRVLEPLG	LVEEGAKTAP	GKKRPVEQSP
AAV_3	QERLQEDTSF	GGNLGRAVFQ	AKKRILEPLG	LVEEAAKTAP	GKKGAVDQSP
AAV_9	QERLQEDTSF	GGNLGRAVFQ	AKKRVLEPLG	LVEEGAKTAP	GKKRPVEQSP
	151				200
AAV_2	.VEPDSSSGT	GKAGQQPARK	RLNFGQTGDA	DSVPDPQPLG	QPPAAPSGLG
AAV_8	QRSPDSSTGI	GKKGQQPARK	RLNFGQTGDS	ESVPDPQPLG	EPPAAPSGVG
AAV_1	.QEPDSSSGI	GKTGQQPAKK	RLNFGQTGDS	ESVPDPQPLG	EPPATPAAVG
AAV_3	.QEPDSSSGV	GKSGKQPARK	RLNFGQTGDS	ESVPDPQPLG	EPRAAPTSLG
AAV_9	QE.PDSSSGI	GKSGQQPAKK	RLNFGQTGDS	ESVPDPQPLG	EPPEAPSGLG
	201				250
AAV_2	TNTMATGSGA	PMADNNEGAD	GVGNSSGNWH	CDSTWMGDRV	ITTSTRTWAL
AAV_8	PNTMAAGGGA	PMADNNEGAD	GVGSSSGNWH	CDSTWLGDRV	ITTSTRTWAL
AAV_1	PTTMASGGGA	PMADNNEGAD	GVGNASGNWH	CDSTWLGDRV	ITTSTRTWAL
AAV_3	SNTMASGGGA	PMADNNEGAD	GVGNSSGNWH	CDSQWLGDRV	ITTSTRTWAL
AAV_9	PNTMASGGGA	PMADNNEGAD	GVGNSSGNWH	CDSTWLGDRV	ITTSTRTWAL
	251				300
AAV_2	PTYNNHLYKQ	ISSQS--GAS	NDNHYFGYST	PWGYFDENRF	HCHFSPRDWQ
AAV_8	PTYNNHLYKQ	ISNGTSGGAT	NDNTYFGYST	PWGYFDENRF	HCHFSPRDWQ
AAV_1	PTYNNHLYKQ	ISSAST.GAS	NDNHYFGYST	PWGYFDENRF	HCHFSPRDWQ
AAV_3	PTYNNHLYKQ	ISSQS..GAS	NDNHYFGYST	PWGYFDENRF	HCHFSPRDWQ
AAV_9	PTYNNHLYKQ	ISNGTSGGST	NDNTYFGYST	PWGYFDENRF	HCHFSPRDWQ

Fig. 2B

301 350
 AAV_2 RLINNNWGFR PKRLNFKLFN IQVKEVTQND GTTTIANNLT STVQVFTDSE
 AAV_8 RLINNNWGFR PKRLSFKLFN IQVKEVTQNE GTKTIANNLT STIQVFTDSE
 AAV_1 RLINNNWGFR PKRLNFKLFN IQVKEVTTND GVTTIANNLT STVQVFSOSE
 AAV_3 RLINNNWGFR PKKLSFKLFN IQVRGVTQND GTTTIANNLT STVQVFTDSE
 AAV_9 RLINNNWGFR PKRLNFKLFN IQVKEVTTNE GTKTIANNLT STVQVFTDSE

351 400
 AAV_2 YQLPYVLGSA HQGCLPPFPA DVFMVPQYGY LTLNNGSQAV GRSSFYCLEY
 AAV_8 YQLPYVLGSA HQGCLPPFPA DVFMIPQYGY LTLNNGSQAV GRSSFYCLEY
 AAV_1 YQLPYVLGSA HQGCLPPFPA DVFMIPQYGY LTLNNGSQAV GRSSFYCLEY
 AAV_3 YQLPYVLGSA HQGCLPPFPA DVFMVPQYGY LTLNNGSQAV GRSSFYCLEY
 AAV_9 YQLPYVLGSA HQGCLPPFPA DVFMVPQYGY LTLNNGSQAL GRSSFYCLEY

401 450
 AAV_2 FPSQMLRTGN NTFESYTFED VPFHSSYAHS QSLDRLMNPL IDQYLYYLSR
 AAV_8 FPSQMLRTGN NFQFTYTFED VPFHSSYAHS QSLDRLMNPL IDQYLYYLSR
 AAV_1 FPSQMLRTGN NTFESYTFEE VPFHSSYAHS QSLDRLMNPL IDQYLYYLSR
 AAV_3 FPSQMLRTGN NFQFSYTFED VPFHSSYAHS QSLDRLMNPL IDQYLYYLSR
 AAV_9 FPSQMLRTGN NFQFSYTFED VPFHSSYAHS QSLDRLMNPL IDQYLYYLSR

451 500
 AAV_2 TNPSTG.TTT QSRLQFSQAG ASDIRDQS RNWLPGPCYRQQ RVSKTSADNN
 AAV_8 TQTTGG.TAN TQTLGFSQGG PNTMANQA KNWLPGPCYRQQ RVSTTTGQNN
 AAV_1 TQ.NQSGSAQ NKDLLFSRGS PAGMSVQP KNWLPGPCYRQQ RVSKTKTDNN
 AAV_3 TQTTSGTTN QSRLQFSQAG PQSMSLQA RNWLPGPCYRQQ RLSKTANDNN
 AAV_9 TQTTG..TGG TQTLAFSQAG PSSMANQA RNWLPGPCYRQQ RVSTTTNQNN

501 550
 AAV_2 NSEYSWTGAT KYHLNGRDSL VNPGPAMASH KDDEEKFFPQ SGVLIFGKQG
 AAV_8 NSNFAWTAGT KYHLNGRNSL ANPGIAMATH KDDEERFFPS NGILIFGKQN
 AAV_1 NSNFTWTGAS KYNLNGRESI INPGTAMASH KDDEDKFFPM SGVMIFGKES
 AAV_3 NSNFPWTAAS KYHLNGRDSL VNPGPAMASH KDDEEKFFPM HGNLIFGKEG
 AAV_9 NSNFAWTGAA KFKLNGRDSL MNPGVAMASH KDDEDRFFPS SGVLIFGKQG

551 600
 AAV_2 SEKTNVNDIEK VMITDEEEIR T TNPVATEQY GSVSTNLQRG NRQAATADVN
 AAV_8 AARDNADYSO VMLTSEEEIK T TNPVATEEY GIVADNLQQO NTAPQIGTVN
 AAV_1 AGASNTALDN VMITDEEEIK A TNPVATERF GTVAVNFQSS STDPATGDVH
 AAV_3 TTASNAELDN VMITDEEEIR T TNPVATEQY GTVANNLQSS NTAPTGTGVN
 AAV_9 AGNDGVDYSO VLITDEEEIK A TNPVATEEY GAVAINNQAA NTQAQTGLVH

Fig. 2C

	601		650
AAV_2	TQGVLP	QDRDVYLQGP	IWAKIPHTDG HFHPSPLMGG FGLKHPPPQI
AAV_8	SQGALP	QNRDVYLQGP	IWAKIPHTDG NFHPSPLMGG FGLKHPPPQI
AAV_1	AMGALP	QDRDVYLQGP	IWAKIPHTDG HFHPSPLMGG FGLKNPPPQI
AAV_3	HQGALP	QDRDVYLQGP	IWAKIPHTDG HFHPSPLMGG FGLKHPPPQI
AAV_9	NQGVIP	QNRDVYLQGP	IWAKIPHTDG NFHPSPLMGG FGLKHPPPQI

	651		700
AAV_2	LIKNTPVPA	NPSTTFSAAKF	ASFITQYSTG QVSVEIEWEL QKENSkrwnp
AAV_8	LIKNTPVPA	DPPTTFNQSKL	NSFITQYSTG QVSVEIEWEL QKENSkrwnp
AAV_1	LIKNTPVPA	NPPAEFSATKF	ASFITQYSTG QVSVEIEWEL QKENSkrwnp
AAV_3	MIKNTPVPA	NPPTTFSPAKF	ASFITQYSTG QVSVEIEWEL QKENSkrwnp
AAV_9	LIKNTPVPA	DPPLTFNQAKL	NSFITQYSTG QVSVEIEWEL QKENSkrwnp

	701		739
AAV_2	EIQYTSNYNK	SVNVDFTVDT	NGVYSEPRPI GTRYLTRNL
AAV_8	EIQYTSNYK	STSVDFAVNT	EGVYSEPRPI GTRYLTRNL
AAV_1	EVQYTSNYAK	SANVDFTVDN	NGLYTEPRPI GTRYLTRPL
AAV_3	EIQYTSNYNK	SVNVDFTVDT	NGVYSEPRPI GTRYLTRNL
AAV_9	EIQYTSNYK	STNVDFAVNT	EGVYSEPRPI GTRYLTRNL

Fig. 3A

Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
 1 5 10 15

Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
 20 25 30

Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Arg Asn Leu Ile
 35 40 45

Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
 50 55 60

Val Gln Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
 65 70 75 80

Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Leu His Val Leu Val Glu
 85 90 95

Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
 100 105 110

Arg Glu Lys Leu Val Gln Thr Ile Tyr Arg Gly Ile Glu Pro Thr Leu
 115 120 125

Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
 130 135 140

Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
 145 150 155 160

Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Glu Tyr Ile
 165 170 175

Ser Ala Cys Leu Asn Leu Ala Glu Arg Lys Arg Leu Val Ala Gln His
 180 185 190

Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Leu Asn
 195 200 205

Fig. 3B

Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
 210 215 220

Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys
 225 230 235 240

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
 245 250 255

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
 260 265 270

Ile Met Ala Leu Thr Lys Ser Ala Pro Asp Tyr Leu Val Gly Pro Ser
 275 280 285

Leu Pro Val Asp Ile Thr Gln Asn Arg Ile Tyr Arg Ile Leu Gln Leu
 290 295 300

Asn Gly Tyr Asp Pro Ala Tyr Ala Gly Ser Val Phe Leu Gly Trp Ala
 305 310 315 320

Gln Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
 325 330 335

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro
 340 345 350

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
 355 360 365

Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
 370 375 380

Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
 385 390 395 400

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
 405 410 415

Fig. 3C

Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
 420 425 430

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
 435 440 445

Glu Leu Thr Arg Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln
 450 455 460

Glu Val Lys Glu Phe Phe Arg Trp Ala Ser Asp His Val Thr Glu Val
 465 470 475 480

Ala His Glu Phe Tyr Val Arg Lys Gly Gly Ala Ser Lys Arg Pro Ala
 485 490 495

Pro Asp Asp Ala Asp Lys Ser Glu Pro Lys Arg Ala Cys Pro Ser Val
 500 505 510

Ala Asp Pro Ser Thr Ser Asp Ala Glu Gly Ala Pro Val Asp Phe Ala
 515 520 525

Asp Arg Tyr Gln Asn Lys Cys Ser Arg His Ala Gly Met Leu Gln Met
 530 535 540

Leu Leu Pro Cys Lys Thr Cys Glu Arg Met Asn Gln Asn Phe Asn Ile
 545 550 555 560

Cys Phe Thr His Gly Val Arg Asp Cys Ser Glu Cys Phe Pro Gly Val
 565 570 575

Ser Glu Ser Gln Pro Val Val Arg Lys Arg Thr Tyr Arg Lys Leu Cys
 580 585 590

Ala Ile His His Leu Leu Gly Arg Ala Pro Glu Ile Ala Cys Ser Ala
 595 600 605

Cys Asp Leu Val Asn Val Asp Leu Asp Asp Cys Val Ser Glu Gln
 610 615 620

SEQUENCE LISTING

<110> The Trustees of The University of Pennsylvania
 Gao, Guangping
 Wilson, James M.
 Alvirà, Mauricio

<120> Adeno-Associated Virus (AAV) Serotype 9 Sequences, Vectors
 Containing Same, and Uses Therefor

<130> UPN-02734PCT

<150> US 60/341,150
 <151> 2001-12-17

<150> US 60/386,132
 <151> 2002-06-05

<160> 7

<170> PatentIn version 3.1

<210> 1
 <211> 4382
 <212> DNA
 <213> adeno-associated virus serotype 9

<400> 1
 cagagaggga gtggccaact ccatacactag gggtaatcgc gaagcgcctc ccacgctgcc 60
 gcgtcagcgc tgacgtagat tacgtcatag gggagtggtc ctgtattagc tgtcacgtga 120
 gtgcttttgc gacattttgc gacaccacat ggccatttga ggtatatatg gccgagtga 180
 cgagcaggat ctccattttg accgcgaaat ttgaacgagc agcagccatg ccgggcttct 240
 acgagattgt gatcaagggtg ccgagcgacc tggacgagca cctgccgggc atttctgact 300
 cttttgtgaa ctgggtggcc gagaaggaat gggagctgcc cccggattct gacatggatc 360
 ggaatctgat cgagcaggca cccctgaccg tggccgagaa gctgtagcgc gacttcctgg 420
 tccaatggcg ccgcgtgagt aaggccccgg aggccctctt ctttgttcag ttcgagaagg 480
 gcgagagcta ctttcacctg cacgttcttg tcgagaccac gggggtcaag tccatggtgc 540
 taggcgctt cctgagtcag attcgggaga agctgggtcca gaccatctac cgcgggatcg 600
 agccgaccct gcccaactgg ttcgcggtga ccaagacgcg taatggcgcc ggcgggggga 660
 acaagggtgt ggacgagtgc tacatcccca actacctct gcccaagact cagcccgagc 720
 tgcagtgggc gtggactaac atggaggagt atataagcgc gtgcttgaac ctggccgagc 780
 gcaaacggct cgtggcgag cacctgaccc acgtcagcca gacgcaggag cagaacaagg 840
 agaatctgaa cccaattct gacgcgcccg tgatcaggtc aaaaacctcc gcgcgctaca 900

tggagctggt cgggtggctg gtggaccggg gcatcacctc cgagaagcag tggatccagg 960
aggaccaggc ctctacatc tccttcaacg ccgcctccaa ctgcggtcc cagatcaagg 1020
ccgcgctgga caatgccggc aagatcatgg cgctgaccaa atccgcgcc gactacctgg 1080
taggcccttc acttccggtg gacattacgc agaaccgcat ctaccgcatc ctgcagctca 1140
acggctacga cctgcctac gccggctccg tctttctcgg ctgggcacaa aagaagttcg 1200
ggaaacgcaa caccatctgg ctgtttgggc cggccaccac gggaaagacc aacatcgag 1260
aagccattgc ccacgccgtg cccttctacg gctgcgtcaa ctggaccaat gagaactttc 1320
ccttcaacga ttgcgtcgac aagatggtga tctggtggga ggagggaag atgacggcca 1380
aggctcgtgga gtccgccaag gccattctcg gcggcagcaa ggtgcgctg gaccaaagt 1440
gcaagtcgtc cgcgcagatc gacccactc ccgtgatcgt cacctccaac accaactgt 1500
gcgccgtgat tgacgggaac agcaccacct tcgagcacca gcagcctctc caggaccgga 1560
tgtttaagtt cgaactcacc cgcgctctgg agcacgactt tggcaagggtg acaaagcagg 1620
aagtcaaaga gttcttccgc tgggccagtg atcacgtgac cgaggtggcg catgagtttt 1680
acgtcagaaa gggcggagcc agcaaaagac ccgccccga tgacgcggat aaaagcgagc 1740
ccaagcgggc ctgcccctca gtgcgggatc catcgacgtc agacgcggaa ggagctccgg 1800
tggactttgc cgacaggtac caaaacaaat gttctcgtca cgcgggcatg cttcagatgc 1860
tgcttccctg caaaacgtgc gagagaatga atcagaattt caacatttgc ttcacacag 1920
gggtcagaga ctgctcagag tgtttccccg gcgtgtcaga atctcaaccg gtcgtcagaa 1980
agaggacgta tcggaaactc tgtgcgattc atcatctgct ggggcgggct cccgagattg 2040
cttgctcggc ctgcgatctg gtcaacgtgg acctggatga ctgtgtttct gagcaataaa 2100
tgacttaaac caggtatggc tgccgatggt tatcttccag attggctcga ggacaacctc 2160
tctgagggca ttcgcgagtg gtgggacctg aaacctggag cccgaaacc caaagccaac 2220
cagcaaaagc aggacgacgg ccgggtctg gtgcttctg gctacaagta cctcggaacc 2280
ttcaacggac tcgacaaggg ggagcccgtc aacgcggcgg acgcagcggc cctcgagcac 2340
gacaaggcct acgaccagca gctcaaagcg ggtgacaatc cgtacctgcg gtataaccac 2400
gccgacgccg agtttcagga gcgtctgcaa gaagatacgt cttttggggg caacctcggg 2460
cgagcagtct tccaggccaa gaagcgggtt ctggaacctc tcggtctggt tgaggaaggc 2520
gctaagacgg ctcttgaaa gaagagaccg gtagagcagt caccccaaga accagactca 2580

tcctcgggca tcggcaaatc aggccagcag cccgctaaaa agagactcaa ttttggtcag	2640
actggcgact cagagtcagt ccccgacca caacctctcg gagaacctcc agaagcccc	2700
tcaggtctgg gacctaatc aatggcttca ggcggtggcg ctccaatggc agacaataac	2760
gaaggcgccg acggagtggg taattcctcg ggaaattggc attgcgattc cacatggctg	2820
ggggacagag tcatcaccac cagcaccga acctgggcat tgcccaccta caacaaccac	2880
ctctacaagc aaatctccaa tggaacatcg ggaggaagca ccaacgaca cacctacttt	2940
ggctacagca cccctgggg gtattttgac ttcaacagat tccactgcca cttctacca	3000
cgtgactggc agcgactcat caacaacaac tggggattcc ggccaaagag actcaacttc	3060
aagctgttca acatccaggt caaggagggt acgacgaacg aaggcaccaa gaccatcgcc	3120
aataacctta ccagcacctg ccagggtctt acggactcgg agtaccagct accgtacgtc	3180
ctaggctctg cccaccaagg atgcctgcca ccgtttcctg cagacgtctt catggttcct	3240
cagtacggct acctgacgct caacaatgga agtcaagcgt taggacgttc ttctttctac	3300
tgtctggaat acttcccttc tcagatgctg agaaccggca acaactttca gttcagctac	3360
actttcgagg acgtgccttt ccacagcagc tacgcacaca gccagagtct agatcgactg	3420
atgaaccccc tcatcgacca gtacctatac tacctgggtca gaacacagac aactggaact	3480
gggggaactc aaactttggc attcagccaa gcaggcccta gctcaatggc caatcaggct	3540
agaaactggg taccggggcc ttgctaccgt cagcagcgcg tctccacaac caccaacca	3600
aataacaaca gcaactttgc gtggacggga gctgctaaat tcaagctgaa cgggagagac	3660
tcgctaata atcctggcgt ggctatggca tcgcacaaag acgacgagga ccgcttcttt	3720
ccatcaagtg gcgttctcat atttggcaag caaggagccg ggaacgatgg agtcgactac	3780
agccagggtc tgattacaga tgaggaagaa attaaagcca ccaaccctgt agccacagag	3840
gaatacggag cagtggccat caacaaccag gccgctaaca cgcaggcgca aactggactt	3900
gtgcataacc agggagttat tcctgggtatg gtctggcaga accgggacgt gtacctgcag	3960
ggcctatatt gggctaaaat acctcacaca gatggcaact ttacccgtc tcctctgatg	4020
ggtggatttg gactgaaaca cccacctcca cagattctaa ttaaaaatac accagtgccg	4080
gcagatcctc ctcttacctt caatcaagcc aagctgaact ctttcatcac gcagtacagc	4140
acgggacaag tcagcgtgga aatcgagtgg gagctgcaga aagaaaacag caagcgctgg	4200
aatccagaga tccagtatac ttcaaactac taaaatcta caaatgtgga ctttgctgtc	4260
aataccgaag gtgtttactc tgagcctcgc ccatttggtc ctggttacct caccgtaat	4320

ttgtaattgc ctgttaatca ataaaccggt taattcggtt cagttgaact ttggtctctg 4380
cg 4382

<210> 2
<211> 736
<212> PRT
<213> capsid protein of adeno-associated virus serotype 9

<400> 2

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
145 150 155 160

Lys Ser Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
 180 185 190

Glu Ala Pro Ser Gly Leu Gly Pro Asn Thr Met Ala Ser Gly Gly Gly
 195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
 210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile
 225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
 245 250 255

Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn
 260 265 270

Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
 275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
 290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
 305 310 315 320

Gln Val Lys Glu Val Thr Thr Asn Glu Gly Thr Lys Thr Ile Ala Asn
 325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu
 340 345 350

Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro
 355 360 365

Ala Asp Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn
 370 375 380

Gly Ser Gln Ala Leu Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
 385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr

405	410	415
Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu 420 425 430		
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Val 435 440 445		
Arg Thr Gln Thr Thr Gly Thr Gly Gly Thr Gln Thr Leu Ala Phe Ser 450 455 460		
Gln Ala Gly Pro Ser Ser Met Ala Asn Gln Ala Arg Asn Trp Val Pro 465 470 475 480		
Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Asn Gln Asn 485 490 495		
Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Ala Lys Phe Lys Leu Asn 500 505 510		
Gly Arg Asp Ser Leu Met Asn Pro Gly Val Ala Met Ala Ser His Lys 515 520 525		
Asp Asp Glu Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe Gly 530 535 540		
Lys Gln Gly Ala Gly Asn Asp Gly Val Asp Tyr Ser Gln Val Leu Ile 545 550 555 560		
Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Glu 565 570 575		
Tyr Gly Ala Val Ala Ile Asn Asn Gln Ala Ala Asn Thr Gln Ala Gln 580 585 590		
Thr Gly Leu Val His Asn Gln Gly Val Ile Pro Gly Met Val Trp Gln 595 600 605		
Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His 610 615 620		
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu 625 630 635 640		

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Leu Thr Phe Asn Gln Ala Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 3

<211> 623

<212> PRT

<213> rep protein of adeno-associated virus serotype 9

<400> 3

Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
1 5 10 15

Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
20 25 30

Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Arg Asn Leu Ile
35 40 45

Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
50 55 60

Val Gln Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
65 70 75 80

Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Leu His Val Leu Val Glu
85 90 95

Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
 100 105 110

Arg Glu Lys Leu Val Gln Thr Ile Tyr Arg Gly Ile Glu Pro Thr Leu
 115 120 125

Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
 130 135 140

Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
 145 150 155 160

Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Glu Tyr Ile
 165 170 175

Ser Ala Cys Leu Asn Leu Ala Glu Arg Lys Arg Leu Val Ala Gln His
 180 185 190

Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Leu Asn
 195 200 205

Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
 210 215 220

Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys
 225 230 235 240

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
 245 250 255

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
 260 265 270

Ile Met Ala Leu Thr Lys Ser Ala Pro Asp Tyr Leu Val Gly Pro Ser
 275 280 285

Leu Pro Val Asp Ile Thr Gln Asn Arg Ile Tyr Arg Ile Leu Gln Leu
 290 295 300

Asn Gly Tyr Asp Pro Ala Tyr Ala Gly Ser Val Phe Leu Gly Trp Ala
 305 310 315 320

Gln Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala

325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Glu Phe Phe Arg Trp Ala Ser Asp His Val Thr Glu Val		
465	470	475
Ala His Glu Phe Tyr Val Arg Lys Gly Gly Ala Ser Lys Arg Pro Ala		
485	490	495
Pro Asp Asp Ala Asp Lys Ser Glu Pro Lys Arg Ala Cys Pro Ser Val		
500	505	510
Ala Asp Pro Ser Thr Ser Asp Ala Glu Gly Ala Pro Val Asp Phe Ala		
515	520	525
Asp Arg Tyr Gln Asn Lys Cys Ser Arg His Ala Gly Met Leu Gln Met		
530	535	540
Leu Leu Pro Cys Lys Thr Cys Glu Arg Met Asn Gln Asn Phe Asn Ile		
545	550	555
		560

Cys Phe Thr His Gly Val Arg Asp Cys Ser Glu Cys Phe Pro Gly Val
 565 570 575

Ser Glu Ser Gln Pro Val Val Arg Lys Arg Thr Tyr Arg Lys Leu Cys
 580 585 590

Ala Ile His His Leu Leu Gly Arg Ala Pro Glu Ile Ala Cys Ser Ala
 595 600 605

Cys Asp Leu Val Asn Val Asp Leu Asp Asp Cys Val Ser Glu Gln
 610 615 620

<210> 4
 <211> 735
 <212> PRT
 <213> adeno-associated virus serotype 2

<400> 4

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
 1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
 20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
 35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
 85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
 100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
 115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
 145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
 165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
 180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
 195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
 210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
 225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
 245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
 260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
 275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
 290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
 305 310 315 320

Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
 325 330 335

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
 340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp

355		360		365
Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser				
370		375		380
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser				
385		390		395
Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu				
	405		410	415
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg				
	420		425	430
Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr				
	435		440	445
Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln				
	450		455	460
Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly				
465		470		475
Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn				
	485		490	495
Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly				
	500		505	510
Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp				
	515		520	525
Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys				
	530		535	540
Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr				
545		550		555
Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr				
	565		570	575
Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Arg Gln Ala Ala Thr				
	580		585	590

Ala Asp Val Asn Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp
595 600 605

Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr
610 615 620

Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys
625 630 635 640

His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn
645 650 655

Pro Ser Thr Thr Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln
660 665 670

Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys
675 680 685

Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr
690 695 700

Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr
705 710 715 720

Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 5

<211> 736

<212> PRT

<213> adeno-associated virus serotype 1

<400> 5

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
 85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
 100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
 115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
 145 150 155 160

Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
 165 170 175

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
 180 185 190

Ala Thr Pro Ala Ala Val Gly Pro Thr Thr Met Ala Ser Gly Gly Gly
 195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala
 210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile
 225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
 245 250 255

Tyr Lys Gln Ile Ser Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His
 260 265 270

Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe

275	280	285
His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn 290 295 300		
Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln 305 310 315 320		
Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn 325 330 335		
Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro 340 345 350		
Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala 355 360 365		
Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly 370 375 380		
Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro 385 390 395 400		
Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe 405 410 415		
Glu Glu Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp 420 425 430		
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg 435 440 445		
Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn Lys Asp Leu Leu Phe Ser 450 455 460		
Arg Gly Ser Pro Ala Gly Met Ser Val Gln Pro Lys Asn Trp Leu Pro 465 470 475 480		
Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Lys Thr Asp Asn 485 490 495		
Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala Ser Lys Tyr Asn Leu Asn 500 505 510		

Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His Lys
 515 520 525

Asp Asp Glu Asp Lys Phe Phe Pro Met Ser Gly Val Met Ile Phe Gly
 530 535 540

Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala Leu Asp Asn Val Met Ile
 545 550 555 560

Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Arg
 565 570 575

Phe Gly Thr Val Ala Val Asn Phe Gln Ser Ser Ser Thr Asp Pro Ala
 580 585 590

Thr Gly Asp Val His Ala Met Gly Ala Leu Pro Gly Met Val Trp Gln
 595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
 610 615 620

Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
 625 630 635 640

Lys Asn Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
 645 650 655

Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys Phe Ala Ser Phe Ile Thr
 660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
 675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Val Gln Tyr Thr Ser Asn
 690 695 700

Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr Val Asp Asn Asn Gly Leu
 705 710 715 720

Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Pro Leu
 725 730 735

<210> 6
 <211> 736
 <212> PRT
 <213> adeno-associated virus serotype 3

 <400> 6

 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
 1 5 10 15

 Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Val Pro Gln Pro
 20 25 30

 Lys Ala Asn Gln Gln His Gln Asp Asn Arg Arg Gly Leu Val Leu Pro
 35 40 45

 Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60

 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80

 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
 85 90 95

 Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
 100 105 110

 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Ile Leu Glu Pro
 115 120 125

 Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Gly
 130 135 140

 Ala Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly
 145 150 155 160

 Lys Ser Gly Lys Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
 165 170 175

 Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
 180 185 190

 Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly

195	200	205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser 210 215 220		
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile 225 230 235 240		
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu 245 250 255		
Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr 260 265 270		
Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His 275 280 285		
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp 290 295 300		
Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val 305 310 315 320		
Arg Gly Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu 325 330 335		
Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr 340 345 350		
Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp 355 360 365		
Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser 370 375 380		
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser 385 390 395 400		
Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe Glu 405 410 415		
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg 420 425 430		

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg Thr
 435 440 445

Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe Ser
 450 455 460

Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro
 465 470 475 480

Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp Asn
 485 490 495

Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu Asn
 500 505 510

Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys
 515 520 525

Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe Gly
 530 535 540

Lys Glu Gly Thr Thr Ala Ser Asn Ala Glu Leu Asp Asn Val Met Ile
 545 550 555 560

Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln
 565 570 575

Tyr Gly Thr Val Ala Asn Asn Leu Gln Ser Ser Asn Thr Ala Pro Thr
 580 585 590

Thr Gly Thr Val Asn His Gln Gly Ala Leu Pro Gly Met Val Trp Gln
 595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
 610 615 620

Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
 625 630 635 640

Lys His Pro Pro Pro Gln Ile Met Ile Lys Asn Thr Pro Val Pro Ala
 645 650 655

Asn Pro Pro Thr Thr Phe Ser Pro Ala Lys Phe Ala Ser Phe Ile Thr
 660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
 675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
 690 695 700

Tyr Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val
 705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
 725 730 735

<210> 7
 <211> 738
 <212> PRT
 <213> adeno-associated virus serotype 8
 <400> 7

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
 1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Lys Pro
 20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
 35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80

Gln Gln Leu Gln Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
 85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
 100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro

115					120					125					
Leu	Gly	Leu	Val	Glu	Glu	Gly	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg
130						135					140				
Pro	Val	Glu	Pro	Ser	Pro	Gln	Arg	Ser	Pro	Asp	Ser	Ser	Thr	Gly	Ile
145					150					155					160
Gly	Lys	Lys	Gly	Gln	Gln	Pro	Ala	Arg	Lys	Arg	Leu	Asn	Phe	Gly	Gln
				165					170					175	
Thr	Gly	Asp	Ser	Glu	Ser	Val	Pro	Asp	Pro	Gln	Pro	Leu	Gly	Glu	Pro
			180					185					190		
Pro	Ala	Ala	Pro	Ser	Gly	Val	Gly	Pro	Asn	Thr	Met	Ala	Ala	Gly	Gly
	195						200					205			
Gly	Ala	Pro	Met	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser
	210					215					220				
Ser	Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Thr	Trp	Leu	Gly	Asp	Arg	Val
225					230					235					240
Ile	Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His
				245					250					255	
Leu	Tyr	Lys	Gln	Ile	Ser	Asn	Gly	Thr	Ser	Gly	Gly	Ala	Thr	Asn	Asp
			260					265					270		
Asn	Thr	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn
		275					280					285			
Arg	Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn
	290					295					300				
Asn	Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Ser	Phe	Lys	Leu	Phe	Asn
305					310					315					320
Ile	Gln	Val	Lys	Glu	Val	Thr	Gln	Asn	Glu	Gly	Thr	Lys	Thr	Ile	Ala
				325					330					335	
Asn	Asn	Leu	Thr	Ser	Thr	Ile	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln
			340					345					350		

Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
 355 360 365

Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn.
 370 375 380

Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
 385 390 395 400

Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Thr Tyr
 405 410 415

Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
 420 425 430

Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
 435 440 445

Ser Arg Thr Gln Thr Thr Gly Gly Thr Ala Asn Thr Gln Thr Leu Gly
 450 455 460

Phe Ser Gln Gly Gly Pro Asn Thr Met Ala Asn Gln Ala Lys Asn Trp
 465 470 475 480

Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Gly
 485 490 495

Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Ala Gly Thr Lys Tyr His
 500 505 510

Leu Asn Gly Arg Asn Ser Leu Ala Asn Pro Gly Ile Ala Met Ala Thr
 515 520 525

His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Asn Gly Ile Leu Ile
 530 535 540

Phe Gly Lys Gln Asn Ala Ala Arg Asp Asn Ala Asp Tyr Ser Asp Val
 545 550 555 560.

Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
 565 570 575

Glu Glu Tyr Gly Ile Val Ala Asp Asn Leu Gln Gln Gln Asn Thr Ala
580 585 590

Pro Gln Ile Gly Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
595 600 605

Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
610 615 620

Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625 630 635 640

Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val
645 650 655

Pro Ala Asp Pro Pro Thr Thr Phe Asn Gln Ser Lys Leu Asn Ser Phe
660 665 670

Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
675 680 685

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
690 695 700

Ser Asn Tyr Tyr Lys Ser Thr Ser Val Asp Phe Ala Val Asn Thr Glu
705 710 715 720

Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
725 730 735

Asn Leu

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.